THE

LONDON, EDINBURGH, AND DUBLIN

PHILOSOPHICAL MAGAZINE

AND

JOURNAL OF SCIENCE.

[SEVENTH SERIES.]

JANUARY 1937.

I. *Sparking Potentials at Ultra-High Frequencies.* By JOHN THOMSON, *M.A., D.Sc., Natural Philosophy Department, the University, Glasgow* *.

I. *Introduction.*

THE object of this paper is to extend the range of previous observations [1] on high-frequency sparking-potentials at low gas-pressures to what is now called the "ultra-short wave" region, and to find how far the general results obtained in the previous investigations apply therein. It is also the purpose of the communication to correct an arithmetical error which appeared in the earlier paper, and to call attention to the fact that this correction invalidates a conclusion already drawn by the writer. In the present experiments much greater care has been exercised in securing the accuracy of the measuring apparatus, and although there may still be a faint doubt as to the reliability of certain measurements at the highest frequencies, all the tests known to the writer have been made and are described in detail. The general conclusion to be drawn from the results of the present investigation is that while in certain respects and over a certain range of frequency the sparking-potentials of a hydrogen-filled discharge-tube are given

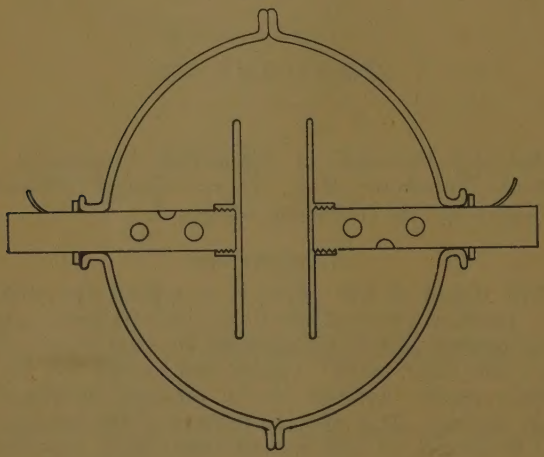
* Communicated by the Author.

by a very simple relation, yet the mechanism by which the discharge is initiated is by no means simple, and involves both electrodic and gaseous phenomena.

II. *Description of Apparatus and Methods of Measurement.*

The Discharge-Tubes.—For reasons given in the previous communication [1], the distance between the parallel disk electrodes in the tubes used in the present experiments was made much smaller than before. In tubes

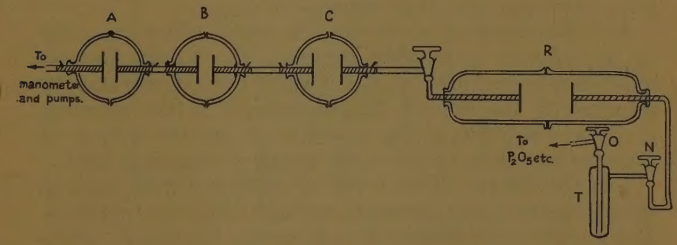
Fig. 1.



A and B (fig. 2) the inter-electrode distance (d) was 2·58 cm., while in tube C it was 4·9 cm. In order that the electrostatic field between the electrodes should be as uniform as possible, the diameter chosen for the copper disks was 8 cm. It was found that the tubes to contain such a system had to be made in a special mould, if the other conditions necessary for accurate measurements and freedom from wall effects were to be observed. The final form of tube A is shown in fig. 1, which is drawn to scale. Tube C, which was only used occasionally, was designed to test the mode of variation of the sparking-potential (V_s) with d. The electrode leads in all three

tubes were of copper tubing 1·4 cm. in diameter, cut as short as possible and provided with copper tags at their ends for electrical connexions. Tubes A and B were geometrically and electrically identical within the limits of design. They were built up in the following manner :— First one electrode was sealed in position in one half of the tube by means of a good vacuum cement. Then the other half of the tube, containing the other loose electrode, was placed in position, the two electrodes being separated by an accurately turned parallel-sided block of ebonite of the requisite thickness and pierced with a central hole. The loose electrode was then centred and brought accurately against one face of the block, the other face of which rested flat on the other fixed electrode.

Fig. 2.



In this position the free electrode was cemented to the half-tube, which was then separated from the other half and the ebonite block removed. The two halves were then cemented together. It will be recognized that with glass discharge-tubes this technique offers the greatest possible facility for accurate assembly of the electrodes, and where the half-tubes have their edges well ground there is no difficulty in making good permanent joints.

The Gas System.—The three-stage oil diffusion pump was connected, as in the previous investigation, *via* an oil manometer to the discharge-tubes A, B, C in the order shown in fig. 2. Tube C was joined to a large cylindrical reservoir R (tube B of the previous paper [1]) with a tap between them. At the other end of R the gas purification system and electrolysis apparatus were attached, the former consisting of two tubes of calcium chloride and a tube of phosphorus pentoxide, with, for final purification,

a trap which could be surrounded with liquid air. The electrolyte used was barium hydroxide. The procedure in allowing the gas into the reservoir and tubes was as follows :—The tap O was closed and the trap T was immersed in liquid air. Then for a period of two hours, during which the tubes and reservoir were being pumped free from the last sample of gas, the hydrogen in T was dried. At the end of that time the tap N was opened and enough gas was allowed into the tubes to give the required pressure. The tap N was then closed, the liquid air removed, and the tap O was opened, to remain open until the next sample of gas was required. In this way a very thorough drying of the gas was achieved, and at the same time excess water-vapour and (possibly) carbon dioxide were removed from the trap. As will be shown later, there is good reason to believe that the present system gave purer gas than that used previously.

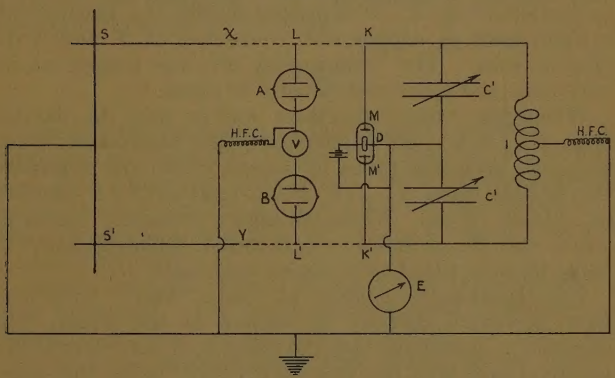
The Measurement of the Gas Pressure.—This was done in the manner previously reported [1]. In the present experiments the tubes A, B, C formed one volume and the reservoir R the other. The constant k in the equation $p = k^n p_0$ was 0·766, giving a very much greater number of observations than before in the range of pressures used.

The Oscillators.—Two types of oscillator were used in these experiments to generate the high-frequency currents. For wave-lengths greater than 4 metres the usual " push-pull " circuit shown in fig. 2 of the previous communication was found to be satisfactory, with certain modifications at the higher frequencies. For example, it was found to be advantageous to insert H.F. chokes in the anode, filament, and grid-potential supply leads, and in certain cases the grid circuit was tuned to the anode circuit by means of a condenser across the grid coil. For wave-lengths less than 4 metres a split-anode magnetron valve with the appropriate Lecher-wire anode circuit was employed. The valve used was the Marconi-Osram S.W. 11.

The Measurement of V_s.—In all cases the oscillators were coupled loosely and electromagnetically to a tuned circuit containing the discharge-tube or tubes, the exact nature of the circuit depending upon the frequency. Fig. 3 is a diagram of a typical arrangement of these coupled circuits, and will be used in explaining the means by which V_s was determined.

Throughout the investigation readings of V_s were taken in each sample of gas at each gas pressure at two different frequencies, one of which was kept constant during all the experiments. This " standard " frequency (about 6·6 megacycles per second) supplied comparison sets of observations by which the behaviour of the discharge-tubes could be checked. Such a test was found to be more reliable than that made in the previous investigation, where a steady source of potential was used for the purpose, for the writer has recently shown [2] that the sparking potential in a discharge-tube, using a steady source of potential, varies according to the fre-

Fig. 3.



quency of the source of potential which has produced the previous discharge. The " test " oscillator was loosely coupled to the circuit shown on the right in fig. 3. The coupling inductance I was shunted by the two equal, variable, transmitting condensers C′, each of which had a maximum capacity of 100 $\mu\mu$F. In parallel with these condensers were the two tubes A and B in series and a double diode valve D. The filament of the latter was connected to the lead between the condensers C′ and to one terminal of a standard electrostatic voltmeter E, the other terminal of which was earthed. The tubes A and B were connected together electrically by means of a vacuojunction V (separate heater type), the gas connexion being a short piece of glass tubing between the

copper leads. To reduce stray capacities the vacuo-junction was dismounted from its four-pin base, and the leads from the heater wire were soldered directly to the tags on the electrode leads of the discharge-tubes. The leads from the thermojunction were connected through high-frequency chokes to a Unipivot microamperemeter. One side of the heater wire was connected through another choke to earth, and a centre-tapping on the coil I was similarly treated. The tuning condensers C′ were varied in such a manner that their capacities were always equal, and so the lead between them, in common with the lead to the vacuojunction V and the centre-tapping of I, was a potential node.

The coupled circuit used at the high frequencies is shown on the left of fig. 3. It consisted simply of a Lecher-wire system made of copper rod, connected at X and Y to flexible wire. The tuning wire SS′ was tapped at its mid-point and connected directly to earth.

When the " test " oscillator was in use, the flexible leads KL, K′L′ were connected to the tags on the leads from the discharge-tubes by means of slip-on terminals, the flexible leads XL, YL′ being removed to a distance of at least 10 cm. When the higher frequency oscillator was in use, the procedure was reversed, the leads XL, YL′ being likewise fitted with slip-on terminals.

The sparking-potentials at the " test " frequency were measured directly by means of the circuit just described. Assuming that the tubes A, B were electrically identical and that the two capacities C′ were the same, then the potential difference indicated on the voltmeter E was the peak potential difference across each tube. Even if the two halves of the diode had different capacities, the ballasting effect of the capacities C′ eliminated the error which might have been produced. But this method of measuring V_s is only accurate at frequencies such that the inductance of the leads KL, KM, K′L′, K′M′ may be neglected, and so, at the higher frequencies, the vacuo-junction V was employed. Using either of the circuits shown in fig. 3, the current in the heater-wire of the vacuo-junction V was proportional to the R.M.S. potential difference across the two tubes, and, assuming the latter to be identical, proportional to the potential difference across either. Assuming for the moment that the " peak factor " was constant for all the experiments made, it

follows that the peak potential difference V across one discharge-tube was related to the current in the heater wire i by the formula

$$V = ki, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where k, of course, varied with the frequency.

A direct determination of k is possible at the frequency of the "test" oscillator. Simultaneous readings were therefore made of the readings of meter E and the thermo-electric current in the vacuojunction, and from these readings V and i were deduced and k calculated. But at this frequency it may be assumed that the circuit LAVBL′ consists only of two equal capacities, for the impedances of the leads and the heater wire in the vacuo-junction are negligible. It was therefore assumed that equation (1) could be rewritten

$$V = ki = \frac{a}{\nu} i,$$

where ν is the frequency of the applied oscillation. The equation was tested at 10 and 2 megacycles per second, and found accurate within the limits of experimental error (about $\frac{1}{2}$ per cent.). The value of a calculated from these experiments was $7{\cdot}93 . 10^7$, where V was in volts and i in milliamperes.

To test the behaviour of the circuit another experiment was performed. The usual condenser and vacuojunction arrangement was set up in parallel with the tube, and simultaneous measurements were taken of the current in the tube vacuojunction and the current in the junction between the two calculable capacities. The latter consisted of small parallel plate condensers with brass disk electrodes about 1 cm. in radius and 0·05 cm. apart, the thickness of the plates being 0·06 cm. Their leads were of brass rod as fine as was consistent with rigidity, and were held firmly in an ebonite frame. These capacities were calculated by the usual formula [3], and hence the absolute value of the impedance of the tube circuit was estimated. The value of a found by this method was $8{\cdot}03 . 10^7$, and although this differs from the first estimate by less than 2 per cent., the value $7{\cdot}93 . 10^7$ was preferred for various reasons.

The good agreement between the two methods provides a check on the "peak factor" of the oscillation. In

the first method V was given directly ; in the second V was assumed to be $\sqrt{2}\,V_{R.M.S.}$. Apparently the precautions taken to keep the coupling loose and the harmonics negligible were satisfactory.

The Measurement of Frequency.—The experiments covered a wave-length range of 170 to 3 metres. From the former value to 15 metres a General Radio Wave-meter, accurate to 1 per cent., was used. For the shorter wave-lengths Lecher-wires were very loosely coupled to the oscillator. The system used finally consisted of copper wires (No. 18 S.W.G.) about 13 cm. apart and about 12 metres long. The sliding connecting piece consisted of a copper disk, 26 cm. in diameter, and pierced above its centroid by two holes to take the wires. Resonance was observed by the reaction on the anode current of the oscillator. Other shorter systems were also used, but the most reliable measurements were made with the long wires. At a wave-length of 5 metres four resonance positions were obtained. A small auxiliary wave-meter (coil and condenser) was also employed to ensure that the frequency of the oscillator did not vary appreciably while it was in use.

III. *Experiments and Results.*

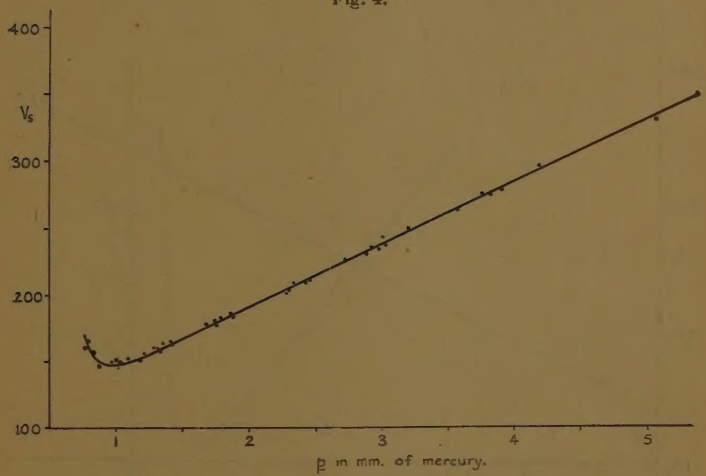
(*a*) As has been mentioned above, along with each set of observations on the variation of V_s with p (the gas pressure) at wave-length λ a comparison set at wave-length λ_0 was taken. The exact value of λ_0 was 47·4 metres, and, on account of the large number of readings obtained at this wave-length, it was possible to deduce some interesting facts concerning high-frequency sparking potentials. To indicate the reliance which may be placed on the purity of the gas used, and, perhaps, on the method of measurement of V_s, all the readings taken at wave-length λ_0 between 5th November, 1935, and 29th January, 1936, are plotted in fig. 4. The abscissa measures p in mm. of mercury and the ordinate V_s in volts, the distance between the electrodes being, of course, 2·58 cm.

The curve representing the mean of these observations has been drawn, and it will be noted that the maximum deviation from this curve, measured parallel to the V_s-axis, is about 5 volts. The meters used to measure V_s were graduated in 2-volt and 5-volt steps, the more sensitive

having a range of 60–240 volts and the less sensitive a range of 200–600 volts. Hence the maximum deviation of an observation from the mean is not much greater than the experimental error, even if p is assumed to be given with a very much greater accuracy.

But this simple analysis does not exhaust the possibilities of the observations. The discrepancies between the readings recorded in fig. 4 are in large measure due to a gradual movement of all the points upwards along the V_s -axis. The greater part of the variations can be

Fig. 4.



explained by a slow increase in A in that part of the curve for which

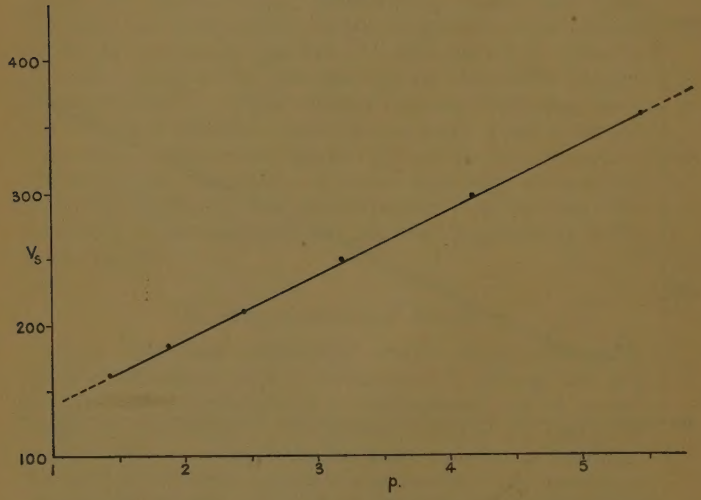
$$V_s = A + Bp. \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

It would appear that during the course of the investigations the value of A increased from about 96 volts to about 102 volts. To illustrate the agreement between observations of *one set*, the curve of fig. 5 has been drawn. The points shown therein were taken in one pressure run in November 1935.

That the value of A in equation (2) may vary even in the course of one experiment, while the value of B remains the same, is illustrated very clearly by the readings shown in fig. 6. These were taken at a frequency of

$1{\cdot}86.10^7$ c.p.s. in September 1935, and the absolute values of V_s are slightly in error, due to imperfection in the measuring circuit. But the relative values, with which we are at the moment concerned, are accurately given. It will be noted that two parallel straight lines may be drawn through two sets of three points, but that neither line passes through the intermediate point. The only reasonable conclusion which may be drawn is that during the operation of taking the value of V_s at

Fig. 5.

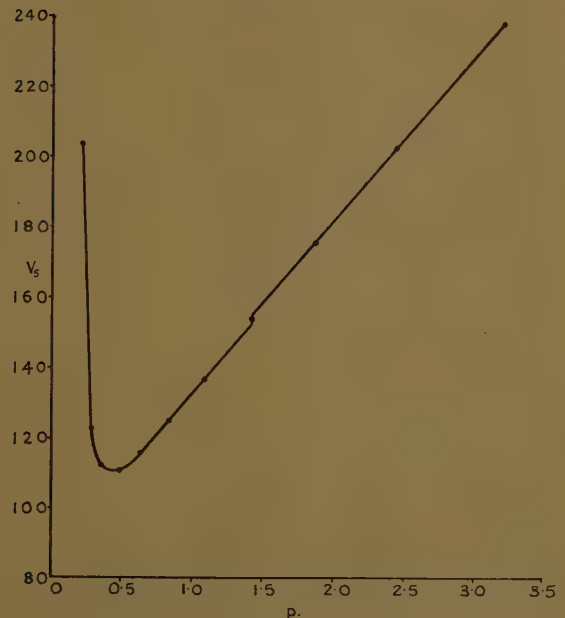


$p=1{\cdot}42$ mm., the value of A for the tube (equation (2)) decreased by about 4 volts. Fig. 6 is not the sole example of this interesting phenomenon. Several curves obtained at different times exhibited the same effect.

(*b*) The V_s against p graphs for various frequencies are shown in fig. 7 (p. 12), the scale being smaller than in the previous figures to allow all the readings to be entered. The range of frequency is considerable—from 1·80 to 99·0 megacycles per second—very much greater than the high-frequency range of the writer's previous experiments. It is at once evident that the conclusion

reached from these earlier experiments cannot be extended to the wider range. In the first place, the double minimum in the lower frequency curves destroys the simplicity of the earlier results, although it may be that the presence of the double minimum is an indication of where a high-frequency theory of the initiation of the discharge might be expected to fail. In the second place,

Fig. 6.



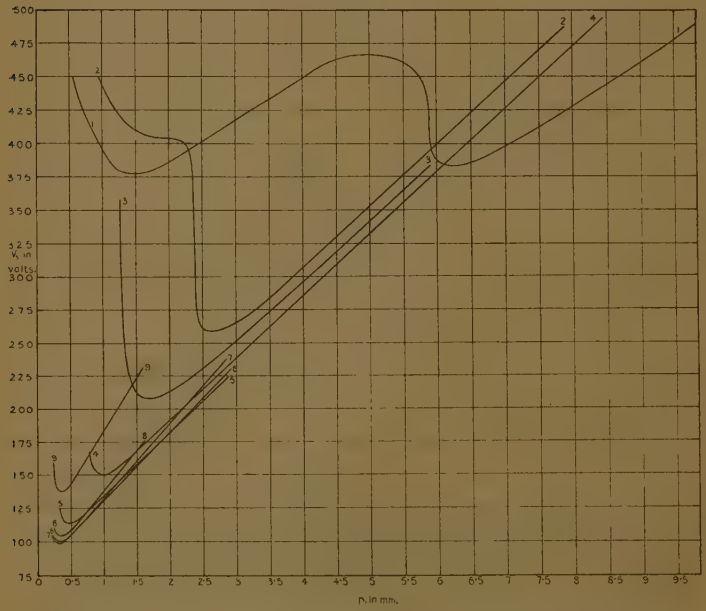
if the present measuring apparatus is to be trusted, the value of B in equation (2), even for those curves which have a single minimum, increases as the frequency increases. Finally, as mentioned in the introduction, the absolute value of B/*d* was erroneously calculated in the earlier paper. The correct value was 26 volts/cm./mm., so that the apparent agreement between B/*d* for the D.C. discharge and the same quantity for the H.F. discharge

vanishes. Table I. of the earlier paper ought, therefore, to read :—

TABLE I.

A	307	325	153	138	107
$\frac{1}{10}$ B	38·8	39	26	26	26
ν	0	851	$8{\cdot}1.10^6$	$1{\cdot}5.10^7$	$2{\cdot}0.10^7$

Fig. 7 *.



The corresponding table for the present experiments, including the minimum value of V_s (V_M) and the pressure p_m, at which the minimum occurs, is :—

A, B/2·58, V_M are given in volts, while p_m is in mm. of mercury. On the other hand, the variation of B with the frequency at the higher values is not so important as the

* See, however, note on p. 24.

table would suggest. In fig. 7 four of the curves for $\nu = 15{\cdot}6$ to $\nu = 23{\cdot}9$ megacycles per second are so close together at high pressures that a slightly larger experimental error would justify superimposing them.

The curve taken at $\nu = 1{\cdot}80$ megacycles per second is not included in the above table. As far as may be judged, the gradient B/2·58 for the high-pressure part of this curve is roughly equal to 18, but the readings begin at too low a value of p to be certain. If a curve is drawn connecting the observed values of p_m and ν, as these are given in Table I. (*a*), the first minimum of the curve at 1·80 megacycles per second, namely, $p_m = 6{\cdot}2$ mm.,

TABLE I. (*a*)*.

No. in fig. 7.	2	3	4	5	6	7	8	9
$10^{-6}\nu$	2·83	3·73	6·33	15·6	49·5	65·4	77·1	99·0
A	121	109	100	87·5	80·0	77	77·6	103
$\frac{1}{2{\cdot}58}$ B	18·0	18·0	18·0	18·5	20·0	21·9	23·9	31·4
V_M	260	209	151	114	100	99	104	138
p_m	2·6	1·7	1·0	0·46	0·315	0·30	0·34	0·36

lies on the natural continuation of this curve. This is shown in fig. 8, where A is the point $\nu = 1{\cdot}80 . 10^6$ c.p.s. and $p_m = 6{\cdot}2$ mm. Incidentally, it should be noted that the results given in this figure are independent of the electrical measurements, and that the curve is very regular.

According to the suggestion made at the end of the earlier paper, B/2·58 of the present investigation should be equal to B/10 of the previous one. That this is not the case must be attributed to a change in the purity of the gas used, for it was again demonstrated in the course of the present work that the relation between V , p, and d for high pressures was of the form

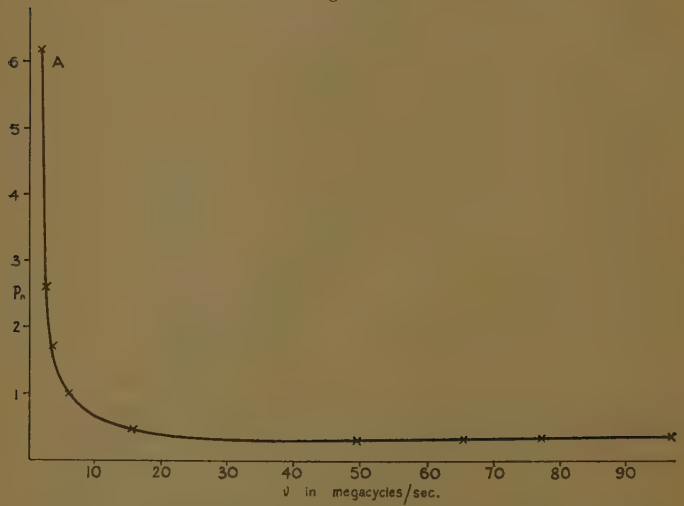
$$V_s = A + bpd.$$

* See note on p. 24.

In verifying this relation, the discharge-tube C was employed in which $d=4{\cdot}9$ cm., and readings were taken at the test frequency. In all cases B/4·9 was 18·0± experimental error. An additional fact which verifies the statement that the gas was of different purity in the two investigations is that the gradient for the initiation of a D.C. discharge in the earlier work was 38·8, while in the present experiment a value of 33·3 was found [2].

The next point arising out of Table I. (*a*) is the fact that the value of $b=\mathrm{B}/2{\cdot}58$, while it is sensibly constant

Fig. 8.



up to a certain frequency, increases fairly rapidly towards 100 megacycles. At the outset the writer was inclined to believe that this increase was evidence of the failure of the measuring apparatus, and he therefore made many tests in the attempt to prove that the circuit already described was not suitable at ultra-high frequencies. The following possible sources of error were investigated :—

(*a*) The inductance of the leads to the electrodes.
(*b*) The absorption by the glass of the tubes.
(*c*) The " skin-effect " in the heater wire of the vacuo-junction.

(*d*) The leakage of high-frequency current through the chokes at the junction of the two discharge-tubes.

Errors (*a*) and (*b*), if they exist at all, produce an apparent *lowering* of the potential difference measured by the vacuojunction method, and so they cannot be held responsible for the apparent *increase* in (*b*). Error (*c*) was estimated from data generously supplied by the Cambridge Scientific Instrument Company, and was found to be less than 1 per cent. at frequencies less than 189 megacycles per second. Error (*d*) was responsible for little variation. The complete elimination of the chokes in the circuit increased (*b*) by a factor of about 5 per cent., but not by more. In the writer's opinion, therefore, the values given in fig. 7 and Table I. (*a*) are correct to 3 per cent., or better even at the highest frequencies *. If, of course, the capacities of the discharge-tubes are functions of the frequency, then the results will be in error by an unknown amount.

IV. *Discussion of the Results.*

At the outset it may be mentioned that the general shapes of the curves of fig. 7 are similar to those found by C. and H. Gutton [4], who used hydrogen gas and external parallel plate electrodes. These investigators employed frequencies ranging from $5{\cdot}33.10^4$ to $9{\cdot}6.10^7$ c.p.s., the distance between the ends of their tube being 10 cm. and the internal diameter 3 cm. The average gradient of the high-pressure parts of their curves, $b=B/10$, for frequencies between $3{\cdot}53.10^6$ and $1{\cdot}25.10^7$ c.p.s. is about 31 volts/cm./mm., but as their electrometer measured "effective" volts, and the figure 31 has been calculated on the assumption that the "peak factor" was $\sqrt{2}$, the value has no accurate significance. On the whole, however, the agreement between the two investigations is satisfactory.

It is now evident that the simple theory of the high-frequency glow, suggested by the writer in 1930 [5] is not completely valid at any of the frequencies so far investigated. At low frequencies (1–5 megacycles per sec.), where the V_s-p curve has two minima, it is probable

* See note on p. 24.

that two different modes of initiation occur. When the pressure is high, the requisite discharge current is brought about by some mechanism which we shall denote by M. When the pressure is low mechanism M is at least complicated, and perhaps entirely superseded, by mechanism N. Double minima of this type have been observed by C. and H. Gutton [4] and by E. W. B. Gill and H. H. Donaldson [6]. The latter, using external electrodes, found that the high-pressure minimum disappeared if the tube walls were away from the discharge. They explained this by saying that the loss of electrons due to their oscillatory motion is greater at low pressures in a short tube than in a long one, and the explanation is supported by the fact that the double minima disappear at higher frequencies, when the amplitude of oscillation of an electron is, of course, smaller. It should be noted, however, that Gill and Donaldson observed a high-pressure minimum at about 8·5 mm. of mercury in air, using an oscillation of wave-length 125 metres. The mean free path of an electron in air at this pressure is only 0·00486 cm., while the distance between the walls of their tube was of the order of 3 cm. One would hardly expect, *a priori*, that the losses would be important under these circumstances. At moderately high frequencies (5–10 megacycles per second) the V_s—p curve is of a relatively simple form, and the straight part is given by the equation

$$V_s = f(\nu) + bpd \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

within the limits of experimental error. In this region the initiation mechanism is very probably that denoted by M, for the M part of the low-frequency curves connects in a natural manner with the moderately high-frequency curve. But at ultra-high frequencies (10–100 megacycles per sec.), b of equation (3) increases as ν increases, and, therefore, in this region also the simple theory (by which b is a gas constant) is invalid. On the other hand, as has already been pointed out, there is a frequency range (15–25 megacycles per second) in which, over the pressures investigated, the sparking potential is *to a first approximation* independent of the frequency at high pressures. The maximum observed difference in V_s at high pressures between 15 and 25 megacycles per second is of the order of 15 volts in 260, and it is on this com-

paratively small difference that the change in gradient has been calculated. It may be well, therefore, even at this stage, to consider a modification of the writer's original hypothesis of the initiation of the discharge, and to ask whether such an hypothesis may not even yet find an application at certain frequencies.

In the earlier theory it was assumed that the typical free electron acted upon by an alternating electric field $E \cos 2\pi\nu t$ began to move at time $t=0$, *i. e.*, when the field-strength was E. Such an electron is not that which will acquire a certain energy most quickly, and the following modified hypothesis is therefore suggested :—

Let the typical free electron begin to move at time $t=0$, when the electric field strength is $E \cos(-\phi)$. Then the ionizing velocity will be most quickly attained if it is acquired in a time t_1 such that the electric field at time t_1 is $E \cos(+\phi)$, for, under these circumstances

$$\int_0^{t_1} \cos(2\pi\nu t - \phi)\, dt$$

is a maximum. Therefore, let $m\ddot{x} = Ee \cos(2\pi\nu t - \phi)$ and $\dot{x}=x=0$ at $t=0$. Let the quantity proportional to the ionization potential of the gas be V_0, and let $\dot{x}=\sqrt{2V_0 e/m}$ at time $t_1 = 2\phi/2\pi\nu$. Also, at $t=t_1$ let $x=L$, where L, as usual, is some distance proportional to the mean free path of the electron.

Then

$$\dot{x} = \frac{E\frac{e}{m}}{2\pi\nu}[\sin\phi + \sin(2\pi\nu t - \phi)] \quad . \quad . \quad . \quad (4)$$

and

$$x = \frac{E\frac{e}{m}}{2\pi\nu}\left[\frac{1}{2\pi\nu}\{\cos\phi - \cos(2\pi\nu t - \phi)\} + t\sin\phi\right]. \quad (5)$$

Therefore

$$\sqrt{2V_0\frac{e}{m}} = \frac{E\frac{e}{m}}{\pi\nu}\sin\phi \quad . \quad . \quad . \quad . \quad (6)$$

and

$$L = \frac{E\frac{e}{m}\phi}{2\pi\nu \cdot \pi\nu}\sin\phi. \quad . \quad . \quad . \quad . \quad . \quad (7)$$

Eliminating ϕ between the two equations,

$$E = \frac{\pi\nu\sqrt{2V_0\frac{e}{m}}}{\frac{e}{m}\sin\left(\frac{2\pi\nu L}{\sqrt{2V_0\frac{e}{m}}}\right)} \quad . \quad . \quad . \quad . \quad . \quad (8)$$

In an infinite uniform alternating field of frequency ν, equation (8) defines the amplitude of the field-strength necessary to produce ionization by electron impact, the thermal energy of the electron being supposed to be small compared with the electron energy. If, however, the velocity of agitation of the typical electron is v_a in the direction of motion, and $\sqrt{2V_0\frac{e}{m}}$ is v_0, equation (8) becomes

$$E = \frac{(v_0 - v_a)\pi\nu}{\frac{e}{m}\sin\left(\frac{2\pi\nu L}{v_0 + v_a}\right)} \quad . \quad . \quad . \quad . \quad . \quad (9)$$

At high pressures L is small, and equation (8) becomes $E = V_0/L$, as in the previous hypothesis, while equation (9) becomes $E = (V_0 - V_A)/L$, where V_A is the energy of agitation of the typical electron.

The writer has recently suggested that the sparking potential V_s of a parallel-plate electrode discharge-tube is made up of two parts, one (A) a function of the electrodes, and possibly of the gas, and the other (b) a function of the gas alone. This appears to be a natural deduction from the fact that at high gas pressures

$$V_s = A + bpd.$$

Let V = the potential difference through which a positive ion must fall to produce a certain number of electrons at the cathode surface, and let V_0, as before, be some quantity proportional to the ionization potential of the gas. Then if $L = K/p$ = the mean free path of an electron in the gas, the fall of potential V_E takes place in a distance $L/4\sqrt{2} = K/4\sqrt{2}p$. Hence at high pressures

$$V_s = V_E + \frac{V_0}{K}p\left(d - \frac{K}{4\sqrt{2}p}\right),$$

combining equation (8) with the above-mentioned hypothesis of space-charge.

Hence

$$V_s = V_E - \left(\frac{V_0}{4\sqrt{2}}\right) + \frac{V_0}{K} pd \quad . \quad . \quad . \quad (10)$$

and $$A = V_E - \frac{V_0}{4\sqrt{2}}, \quad b = \frac{V_0}{K}.$$

The suggestion here is that the production of ions by electron impact in a discharge-tube, before the actual initiation of the self-sustained glow, leads to a condition of space-charge which permits the positive ions to release a sufficient number of electrons from the cathode surface to start the independent current. It is implied in the analysis that, while the space-charge produces an abnormal fall of potential in the immediate neighbourhood of the cathode, the remainder of the inter-electrode distance d possesses a uniform potential gradient. Such conditions are not unlikely, owing to the difference in mobility of the electrons and positive ions. Incidentally it may be mentioned that some form of space-charge initiation of the direct-current glow is strongly suggested by the recent work of Loeb [7] and his collaborators.

To complete the formal analysis, let it be assumed that the same excitation mechanism occurs at lower gas pressures. Then equation (10) takes the general form

$$V_s = V_E + \frac{\pi\nu\sqrt{2V_0\frac{e}{m}}}{\frac{e}{m}\sin\left(\frac{2\pi\nu K}{p\sqrt{2V_0\frac{e}{m}}}\right)}\left\{d - \frac{K}{4\sqrt{2}p}\right\}, \quad (11)$$

and from the experimental results which have been cited it seems probable that V_E is a function of ν.

The function E of equation (8) has a minimum value of $\sqrt{\left\{\frac{2\pi^2\nu^2V_0}{e/m}\right\}}$ at $p_m = \frac{4K\nu}{\sqrt{2V_0 e/m}}$, but the case of V_s in equation (11) is more difficult. The minimum here is given mathematically by the equation

$$\frac{c}{p_m} - \tan\frac{c}{p_m} = \frac{4\sqrt{2}cd}{K}, \quad . \quad . \quad . \quad . \quad (12)$$

where

$$c=\frac{2\pi\nu K}{\sqrt{2V_0e/m}}.$$

But from the nature of the original assumptions (see equations (6) and (7)) c/p_m must be less than $\pi/2$. In this range $\tan x-x$ is always positive, and therefore the equation (12) has no significant root, and the theory has no meaning. It is quite possible, however, that the minimum of E at $p_m=\frac{4K\nu}{\sqrt{2V_0e/m}}$ *is* of significance. Certainly at the highest frequencies employed p_m increases as ν increases.

In the case of the D.C. discharge, when polarization effects have been eliminated, the mode of variation of the sparking-potential with the pressure and electrode distance (the gas being hydrogen) is [2]

$$V_s=320+33{\cdot}3\ pd \quad . \quad . \quad . \quad . \quad (13)$$

at pressures greater than twice the critical value. If V_0 of the above analysis is identified with the ionization potential of the gas and L with the mean free path of an electron in the same gas, then $V_0/K=207$. Therefore, if all non-ionizing collisions are elastic, the electron must travel a distance *in the direction of the electric field* of the order of six times its mean free path before acquiring the ionization potential, since $33{\cdot}3\fallingdotseq 207/6$. At moderately high frequencies, where the amplitude of oscillation of a free electron is still quite large, the value corresponding to 33·3 of equation (13) is 18. This suggests that the electrons in this case travel almost twice as far before producing ionization as in the D.C. discharge-tube. If this interpretation is accepted, it implies that the smaller ionic losses in the H.F. field make it possible for a smaller number of free electrons to excite a discharge, namely, those which preserve their "electrostatic" energy throughout an approximately linear path of the order of 12L, where L is the *mean* free path.

At higher frequencies, however, when the amplitude of oscillation of the electron falls to a value of the same order as L, the undisturbed electronic motion over a long distance is impossible, and, as a consequence, the value of the field strength necessary to produce a discharge must (at low pressures at least) increase. It is therefore

worth while to enquire at what frequency the increase in the field strength (B/d) is to be expected, and this can easily be done.

For the curves taken at frequencies between 6 and 50 megacycles per second, the average value of the electric field strength at 1 mm. pressure may be taken as E=18 volts/cm. The corresponding amplitude of oscillation of a free electron in cm. is $Ee/2\pi\nu^2 m = 5.10^{15}/\nu^2$. Also, at 1 mm. pressure the mean free path of an electron in hydrogen is approximately $7{\cdot}74.10^{-2}$ cm., and 12 times this is of the order of 1 cm. If we put the amplitude of oscillation of the electron equal to 1 cm., then

$$1 = 5.10^{15}/\nu^2$$

or

$$\nu \doteqdot 7.10^7 \text{ cycles/sec.}$$

It is therefore to be expected that the gradient B/d should increase fairly rapidly in the neighbourhood of 70 megacycles per second, and this is in agreement with experiment, as is shown in Table I. (a). It is also to be expected that the gradient will increase very rapidly indeed at frequencies greater than 100 megacycles per second, for the amplitude of oscillation varies inversely as the square of the frequency. It will be difficult to test this point experimentally. In every case the gradient should decrease to a constant value of the order of 18 volts./cm./mm. as the pressure increases. This *can* be tested experimentally with a powerful 3-metre oscillator.

To test the general hypothesis that the electrons producing ionization during the "build-up" of the discharge traverse a distance in the direction of the electric field many times their mean free path it is worth while to examine the critical pressure for the D.C. discharge. In this case it has been suggested above that the ionizing electrons traverse a free path of about 6L. The $V_s - p$ curve might be expected to have a critical point when the pressure is such that 6L is of the same order as d, the inter-electrode distance.

Then

$$p = \frac{6K}{d} = \frac{6.7{\cdot}74.10^{-2}}{2{\cdot}58} \text{ mm.} \doteqdot 0{\cdot}2 \text{ mm.}$$

From experiment the critical pressure is about 0·5 mm., and this is in very good agreement with the theoretical

value, for it is obvious that V_s will begin to increase when 6L is *much less than*, but of the same order as, d. The corresponding test at high frequencies is complicated by the fact that the electron losses are smaller, and that these decrease as the frequency increases. But, since the gradients B/d and the values of p_m, the critical pressure, are of the same order of magnitude, a qualitative agreement is evident in this case also.

V. *Summary.*

1. Experiments are described which exhibit accurately the mode of variation of the sparking-potential of a parallel-plate electrode discharge-tube containing hydrogen with the gas pressure, the inter-electrode distance, and the frequency of the applied alternating potential difference. These experiments lead to the following conclusions :—

(*a*) Using pure dry hydrogen gas in suitable tubes, the sparking-potential for a given pressure, inter-electrode distance, and frequency is very constant. The polarization effects, so troublesome when steady potentials are employed, vanish.

(*b*) At comparatively low frequencies (about 1 megacycle per second) the V_s—p curves have two minima. At higher frequencies (up to 100 megacycles per second) the corresponding curves have only one minimum, and at pressures higher than this minimum value the curves are straight lines. An equation representing the straight parts is :—

$$V_s = f(\nu) + \phi(\nu) pd,$$

where $f(\nu)$ and $\phi(\nu)$ vary very slowly over a certain range of frequencies. When $\phi(\nu)$ is almost constant, its value for the hydrogen used was 18 volts/cm./mm. of mercury.

(*c*) An error appearing in a previous paper by the present writer is corrected, and a deduction made from the erroneous result is retracted.

2. It is shown that the writer's earlier theory of the high-frequency discharge cannot be applied directly to any of the results given above. A modified theory

is then sketched in which the following points are important :—

(*a*) It is suggested that the electric field strength in the region between the parallel-plate electrodes is non-uniform, even before the discharge occurs, and that the inter-electrode distance (d) can be divided into two parts. One of these is comparatively short at high pressures and close to an electrode. In it the fall of potential is large. The other part, approximately equal to d at high pressures, has a uniform potential gradient.

(*b*) It is suggested that the electrons which produce ionization during the "build up" of the discharge traverse approximately linear paths in the direction of the electric field, and that the free paths of such electrons in hydrogen are on the average about 12 times the *mean* free path as calculated from the kinetic theory.

(*c*) By comparison with the results of an earlier paper [2], it is suggested that in the case of a D.C. discharge (in hydrogen) the free path of the ionizing electrons is about six times the kinetic theory's *mean*. The longer paths at higher frequencies are due to the fact that a smaller number of electrons can initiate the discharge, owing to the smaller losses to the electrodes and walls.

(*d*) At very high frequencies the very long ionizing free path is no longer possible, owing to the diminished amplitude of oscillation of a free electron in the electric field. This, it is suggested, explains the increase in $\phi(\nu)$ with the frequency.

(*e*) These suggestions are tested in various ways, and the agreement between theory and experiment is satisfactory.

In conclusion the writer wishes to thank Professor Taylor Jones for the interest he has taken in this investigation, and to express his gratitude to the Carnegie Trust for a grant by means of which some of the more elaborate pieces of apparatus were bought. His thanks are also due to Mr. William Reid for his assistance in mechanical

matters. The experiments were performed in the research laboratories of the Natural Philosophy Department in the University of Glasgow.

References.

(1) Phil. Mag. xviii. p. 696 (1934).
(2) Phil. Mag. xxi. p. 1057 (1936).
(3) Grüneisen and Giebe, *Verh. d. Phys. Ges.* xiv. p. 921 (1912).
(4) *Comptes Rendus*, clxxxvi. p. 303 (1928).
(5) Phil. Mag. x. p. 280 (1930).
(6) Phil. Mag. xii. p. 719 (1931).
(7) Varney, White, Loeb, and Posin, *Phys. Rev.* lxviii. p. 818 (1935).

Note added in the Proof.

Since writing this paper, an error in the vacuojunction readings at very high frequencies has been discovered. This error, which is *not* caused by "skin effect" in the heater wire, is probably due to absorption of the H.F. current in the thermojunction itself. The instrument is so designed that such a phenomenon may readily occur. Columns 8 and 9 of Table I. (*a*) should therefore read as in Table I. (*b*), and the curves 8 and 9 of fig. 7 should be altered to correspond.

TABLE I. (*b*).

	8	9
$10^{-6}\nu$	77·1	99·0
A	75·5	75·0
$\frac{1}{2\cdot58}$B ..	23·3	22·3
V_M	101	98
p_m	0·34	0·36

This correction makes it very uncertain if the apparent variation of B with the frequency is real or not. Experiments designed to test this point thoroughly are now in progress.

II. *On the Energy of Metastable Nitrogen Molecules.*

By Hidenori Hamada *.

1. *Introduction.*

In the afterglow of active nitrogen the selective enhancement of the bands with $v' = \sim 6$ and $v' = \sim 11$ in the first positive band system has been frequently reported. It is also known that the bands with $v' = \sim 17$ are enhanced in the afterglow relatively to those with the neighbouring values of v', though in comparison with the bands above mentioned they appear very faint. A similarly enhanced appearance of the bands is also observed in the band system excited by passing a very weak electric current through the nitrogen at the pressure of a few cm. Hg. The only difference in these two cases is that in the latter the relative intensities of the bands with $v' = \sim 17$ are greater than that observable in the former case. In the case where the nitrogen pressure is low the relative intensities of the bands $v' = \sim 6$ and $v' = \sim 11$ are reversed—that is, the bands with $v' = \sim 6$ appear more intensely than those with $v' = \sim 11$ †.

Assuming a concentrated existence of the metastable non-vibrating A $^3\Sigma$ molecules in the active nitrogen, Cario and Kaplan ‡ concluded that the bands are emitted as the result of a collision of the second kind between these metastable A $^3\Sigma$ molecules and the metastable ^{2}D or ^{2}P atoms, and they thus accounted for the intensity distribution of the bands above-mentioned. On the other hand, Prof. J. Okubo and the present writer § have assumed : (1) In the vibrational states with $v'' = \sim 8$ in the metastable A $^3\Sigma$ state, and also in those $v' = \sim 6$ in the upper B $^3\Pi$ state, the nearer nuclear turning points of the vibration have the same or nearly equal nuclear separations as those in the equilibrium position of the nuclei in the normal molecular state ; (2) with the exception of the neutral unexcited molecules and atoms the metastable A $^3\Sigma^{(\sim 7 \text{ or } \sim 8)}$ molecules are most concentrated in active nitrogen. Further, they have

* Communicated by Prof. J. Okubo.
† J. Okubo and H. Hamada, Astrophys. Journ. lxxvii. p. 130 (1933).
‡ G. Cario and J. Kaplan, *Zeit. f. Phys.* lviii. p. 769 (1929).
§ J. Okubo and H. Hamada, Phys. Rev. xlii. p. 795 (1932).

shown that this intensity distribution follows from the Franck-Condon principle and from the general properties of molecular spectra.

In the case where the molecules are excited by passing an electric current through nitrogen prepared from Kahlbaum's pure sodium azid it was observed that the intensity of the Vegard-Kaplan bands was not sensibly altered, while those of the other bands, caused by a trace of impurity contained in the gas, gradually decreased, this decrease accompanying the progress of the purification by the discharge itself and the cooling with liquid air; and it is therefore believed that the Vegard-Kaplan band system is due to the transition between the state $A\,^3\Sigma$ and $X\,^1\Sigma$ of N_2. Accordingly it follows that the energy of the metastable molecules in active nitrogen, as well as in the nitrogen excited by a weak electric current, is estimated to be 6·14 volts according to the first view and to be about 7·4 volts according to the second. Z. Bay and W. Steiner * have shown that the metastable molecule in active nitrogen excites the mercury atom to the level having the energy of 7·7 volts. E. Lau and O. Reichenheim † have also obtained the same result in the case of the metastable nitrogen molecule in the Geissler tube. These two results seem to agree better with that expected from the second view than that from the first.

It is therefore very important to obtain more precise knowledge regarding the energy of the metastable molecules in active nitrogen and in nitrogen when excited by a very weak electric current. In the following pages the results of the experiment recently undertaken in this direction will be described.

2. *The Band Spectrum of Metastable Nitrogen.*

The nitrogen gas containing the metastable $A\,^3\Sigma$ molecules in a fairly good concentration will for brevity be here called "the metastable nitrogen." This metastable nitrogen differs from the active nitrogen which has the nitrogen atom as a source of energy, and it is possible to prepare these two nitrogens separately, though they are found in combination in some cases of their prepara-

* Z. Bay and W. Steiner, *Zeit. f. Phys. Chemie* (B), ix. p. 93 (1930).
† E. Lau and O. Reichenheim. *Ann. d. Phys.* xii. p. 52 (1932).

tion. Each of them has a characteristic value of lifetime and of energy and emits a characteristic spectrum. As the lifetime of the metastable molecule is long the spectrum of metastable nitrogen is not that of the Vegard-Kaplan band system only.

The tube used for obtaining metastable nitrogen was 3 cm. in internal diameter and the distance between the electrodes was about 10 cm. After the nitrogen gas had been introduced into the tube at a pressure of about a few mm. to a few cm. and was immersed in the liquid air an electric current from a 3·5 kilovolts D.C. dynamo was passed through the gas. By means of a rotating sectored disk the light from the afterglow of the positive column in the interval betweem 0·01 sec. and 0·1 sec. after the cutting off of the discharging current was focussed on the slit of the spectrograph. The bands observed were: (1) Vegard-Kaplan band system; (2) the first positive band system; (3) the second positive band system; and (4) Goldstein-Kaplan band system.

The intensities of all these band systems increase with the increases of the pressure of the gas, but the latter three band systems are relatively enhanced to the first system *, while they decrease with the introduction of other atoms or molecules into the gas in response to the excitation of the spectrum of those introduced. From these results, it may be conceived that when two metastable nitrogen molecules collide with each other it is extremely probable that their energies are employed to excite the three latter band systems mentioned above, and also that the probability of a collision of the second kind taking place between the metastable molecule and the metastable atom which resulted from the collision between two metastable molecules is very small.

An examination of the distribution of the intensity of the bands in the Vegard-Kaplan band system shows that the bands corresponding to smaller vibrational quantum numbers $v'=0$, 1, 2 appear with nearly equal intensity, and therefore that the populations ($\Sigma I^{(v')}/\nu^4$) of molecules in the initial states having these vibrational quantum numbers are nearly equal. As one part of

* When the pressure of nitrogen is very low (10^{-2} mm.) the intensity of the Vegard-Kaplan band system is relatively conspicuous (Kaplan's third nitrogen afterglow, J. Kaplan, Phys. Rev. xlviii. p. 800 (1935)).

the bands corresponding to $v'=3$ or 4 is covered by the superimposition of the second positive band system the amount of populations of these molecules is not definitely known, but it is inferable from the intensity of the members in the shorter wave-length region that these do not much differ from those in the states with $v'=0, 1, 2$.

Owing to the fact that the pre-dissociation takes place at levels higher than that of $v'=6$ or 7 in the A $^3\Sigma$-state the concentration of the molecules in vibrational states with greater quantum numbers cannot be inferred from the intensity of the bands, but this is not the case for states with lower vibrational quantum numbers, and the intensities of the bands having these lower levels as the initial states of the emission will measure the populations of the molecules. From this it is known that the distribution of the molecules in the vibrational states of A $^3\Sigma$ is entirely different from the Boltzmann distribution, and also that the specially high concentration of the molecules in the vibrational state $v'=0$ is not admitted.

In the first positive band system the bands originating from $B^3\Pi(\sim 6)$, $B^3\Pi(\sim 12)$, and $B^3\Pi(\sim 17)$ levels are selectively enhanced, and, similarly, to the case of active nitrogen, the bands from $v'=\sim 12$ are stronger than those from $v'=\sim 6$, but the bands from $v'=\sim 17$, which very weakly appear in active nitrogen, are observed with comparatively great intensity. In the second positive band system the bands from the levels with small values of the vibrational quantum number appear with great intensity.

In the Goldstein-Kaplan band system thirteen bands were observed, namely, 2863·5 (0·2), 3005·4 (0·3), 3159·2 (0·4), 3326·1 (0·5), 3508·0 (0·6), 3707·1 (0·7), 3925·4 (0·8), 4166·0 (0·9) 4432·2 (0·10), 4728·0 (0·11), 5058·6 (0·12), 3025·8 (1·4), and 3178·4 (1·5); but the bands from $v'=1$ were very feeble.

3. *The Energy of Metastable Nitrogen.*

The energy of metastable nitrogen is capable of being estimated from the metallic spectra excited by the metastable nitrogen molecules in the afterglow of the positive column above described, in the case where wire is used as the cathode), as Goldstein * has already observed.

* E. Goldstein, *Phys. Zeit.* vi. p. 14 (1905).

In this experiment the metals used were copper, silver, gold, magnesium, zinc, cadmium, and mercury (in this metal a liquid column about 4 mm. in diameter was

1. Copper.

λ.	I.	V.
5220	2?	6·16
5218		6·16
5153	2?	6·16
4023	2?	6·83
3274	5	3·77
3248	5	3·80
3208	2	5·48
3194	3	5·50
3094	3	5·37
3063	1	5·66
3036	2	5·70
3011	3	5·48
2997	3	5·75
2883	2	5·66
2824	5	5·75
2766	2	6·09
2618	4	6·09
2492	5	4·95
2442	4	5·05
2393	2	6·79
2294	2	6·76
2244	2	5·50
2230	2	6·91
2228		7·17
2215	1	6·95
2200	1	7·24
2200		6·99
2179	1	5·66
2165	1	5·70

2. Silver.

λ.	I.	V.
5472	4	6·01
5466		6·02
5209	3	6·01
4669	3?	6·40
4213	3	6·69
4211		6·69
4055	3?	6·69
3812	2?	6·99
3811		7·00

2. Silver (*con.*).

λ.	I.	V.
3683	2	6·99
3383	5	3·65
3281	5	3·76

3. Gold.

λ.	I.	V.
4812	1?	7·64
4793	2?	7·65
4065	2?	7·64
4041	1?	5·70
3123	4?	5·08
3029	5	5·20
2884	4	6·92
2748	5	5·62
2701	5	5·70
2689	4	7·23
2676	5	4·61
2642	5	5·80
2590	2	7·41
2544	2	7·49
2428	5	5·08
2388	4	6·30
2376	3	6·32
2365	3	7·86
2353	5	6·37
2127	1	5·80

4. Magnesium.

λ.	I.	V.
5184	3	5·09
5173		5·09
5167		5·09
3838	4?	5·92
3832		5·92
3829		5·92
3337	2?	6·40
3332		6·40
3330		6·40

4. Magnesium (*con.*).

λ.	I.	V.
3097	3	6·69
3093		6·69
3091		6·69
2852	5	4·33
2852	1	7·03
2848		7·03
2847		7·03

5. Zinc.

λ.	I.	V.
4811	3	6·62
4722	3?	6·62
4680	3?	6·62
3346–45	2?	7·74
3303	2?	7·74
3076	5	4·01
2139	2	5·77

6. Cadmium.

λ.	I.	V.
5086	1	6·35
4800	2	6·35
4678	2	6·35
3614–11	3	7·34
3468–66	3	7·34
3404	2	7·34
3261	5	3·78
3253	1?	7·72
3081	1	7·72
2982–81	2?	8·06
2837	1	8·06
2288	4	5·39

7. Mercury.

λ.	I.	V.
4358	1	7·69
2537	5	4·86

used as the cathode), and the diameter of these wires was about 1 mm.

The wave-length (λ), the relative intensities (I), and the excitation energies (V in volts) of the lines observed are shown in the table. The lines which are believed to

be definitely excited but are covered by the bands are not described in this table. In this table it may be seen that the spectral terms having excitation potentials smaller than about 7·5 volts are easily excited as the result of the collision of the metastable nitrogen with the metal atoms. The excitation probability of a higher energy term than this value is small, but it may be safely inferred that the highest term which is excited by metastable nitrogen molecule is that of about 8·0 volts.

Evidently in this type of excitation the sharp resonance effect, as in the case of the collision of excited atoms with other atoms, is not to be expected, and the energy of the metastable nitrogen molecules will not be precisely obtained. However, it seems very reasonable to conclude that the metastable nitrogen molecules are distributed among the vibrational levels in the neighbourhood of the level with the energy of about 7·5 volts, where the concentration is at a maximum, instead of assuming that all the metastable nitrogen molecules have a vibrational energy corresponding to the highest initial level of the spectral lines excited by the metastable nitrogen in the spectrograms obtained after an exposure of very long duration. It is known for certain that at all events there is no evidence of a high concentration of the metastable molecules in the non-vibrational state.

4. *The Spectrum of the Afterglow of Active Nitrogen.*

By photographing the afterglow spectrum with a very long exposure and using a sector disk it is found that, besides the first positive band system, there is also a weak appearance of the second positive band system and Vegard-Kaplan band system.

In this first positive band system the bands corresponding to $v' = \sim 6$ and $v' = \sim 11$ are selectively enhanced and the bands with $v' = \sim 17$ are also excited but they are very weak. The Vegard-Kaplan band system, which is intensely excited by a weak current discharge, appears very weakly in this afterglow of the active nitrogen, and therefore a very long time-exposure is required to photograph it. From the comparison of the intensity of this band system and that of the α-bands it seems that the time exposure required to photograph the former system is very long, being about 10^5 times as long as that

required for the latter bands, and this is probably the reason why the former band system was frequently missing in the spectrograms of the afterglow. In this experiment a tube 1 m. in length and 2 cm. in diameter was used, and the discharge was carried out with an induction coil with a condenser of the capacity of $\frac{1}{4}\mu$F in parallel. The light from the afterglow in the tube was observed from the end-on view with a spectrograph modified from the monochromator made by Adam Hilger and Co.

The distribution of the intensity of the bands in the Vegard-Kaplan system in the afterglow was found to be similar to that in the case of the metastable nitrogen. There remains the question of whether the obtained spectrogram is that of the superposition of the spectrum of active nitrogen caused by the atomic nitrogen on those of the second positive band system and of the Vegard-Kaplan system caused by the metastable nitrogen resulting from the discharge. Much was hoped from the observation of the Vegard-Kaplan band system after one second or more after the cutting off of the discharge current, but this was not attempted, as with our apparatus too long an exposure, *i. e.*, more than 1000 hours, would be required. But from the above observations it may safely be concluded that in the active nitrogen the concentration of the metastable molecules with the lower vibrational quantum numbers, and especially those in the non-vibrating state, is negligibly small in comparison with that in the ordinary discharging tube.

5. *On the Nature of Active Nitrogen.*

As above stated, no evidence was obtained from this experiment of the concentrated existence of non-vibrating metastable molecules in active nitrogen. Therefore it seems more appropriate to account for all the observed results in the case of active nitrogen without assuming the concentration of these non-vibrating molecules.

We have formerly shown that the observed results of the afterglow of active nitrogen are followed as the consequence of the Franck-Condon principle by assuming (1) that in the vibrational states with $v'' = \sim 8$, in the metastable $A^3\Sigma$ state, and also in those with $v' = \sim 6$ in the upper $B\,^3\Pi$ state the near nuclear turning points of

the vibration have the same or nearly equal nuclear separations to those in the equilibrium position of nuclei in the normal molecular state ; (2) that with the exception of the neutral unexcited molecules and atoms in active nitrogen the metastable A $^3\Sigma$ molecules in the vibrational states with $v''=7$ or 8 are most concentrated.

The above explanation seems to conflict with the results obtained by H. O. Kneser *, *i. e.*, that the decay of the afterglow of active nitrogen obeys the bimolecular law, but, by taking into consideration the fact that the bands with $v'=\sim 11$ in the α-bands are much intensely enhanced in comparison with the bands with $v'=\sim 6$, together with the fact that the final state (A $^3\Sigma^{(\sim 7)}$) of the emission of the bands with $v'=\sim 11$ has an energy nearly equal to that of two normal atoms, it may be easily understood that their view is consistent with the results Kneser has obtained.

It is necessary again to consider the question whether in the case of the excitation of nitrogen molecules by the triple collision the Franck-Condon principle is applicable or not. Cario † has questioned the application of the principle in the case where two molecules result from a triple collision between two atoms and one molecule. However, in the case of active nitrogen, with the exception of the case of the mixed gases of a small quantity of nitrogen with large quantity of inert gases, the afterglow is believed to be emitted from a nitrogen molecule, and not from the two atoms which recombine by the collision, as G. Herzberg ‡ has stated. Therefore the question is whether the nuclear separation in the equilibrium position of the nitrogen molecule undergoes some change or not in the case where it is excited by collision with a quasi-molecule. But it is considered that the principle holds good for a molecule excited under a proper mechanism of excitation.

On examining the distribution of intensity in the band systems of other molecules excited by active nitrogen, and comparing them with those in cases of excitation caused by electron collision or by white light, it is ascertained that there are two cases, *i. e.*, (1) the distributions are in the same manner in both cases (for example,

* H. O. Kneser, *Ann. d. Phys.* lxxxvii. p. 77 (1928).
† G. Cario, *Zeit. f. Phys.* lxxxix. p. 523 (1934).
‡ G. Herzberg, *Zeit. f. Phys.* xlix. p. 512 (1928).

PN, Se_2, Na_2, I_2, and CO); (2) the distributions are not the same (for example, CN). In the case of excitation by active nitrogen it may be a case of some chemical changes taking place, directly or indirectly, and may also be a case of a manifestation of the effect of collision after the excitation of the molecules. Thus the difference in the intensity distribution in some cases is not sufficient to raise the question as to the applicability of the Franck-Condon principle.

It may be further mentioned here that the accumulation of the metastable molecules in the $A\,^3\Sigma$ ($\sim^7$) state in active nitrogen may also result with greater probability, as Cario has pointed out, from two atoms approaching each other along the repulsive potential curve in comparison with the case where the excitation arises from the normal $X\,^1\Sigma$ state according to the Franck-Condon principle; but it seems preferable to consider that the excitation to the upper $B\,^3\Pi$ ($\sim^6$ and $\sim^{11}$) state from the $A\,^3\Sigma$ ($\sim^7$) state is explained by the Franck-Condon principle.

The present experiment was carried out in the Physical Laboratory of the Tohoku Imperial University, Sendai, under the direction of Prof. J. Okubo, to whom the writer desires to express his sincere thanks.

The Research Laboratory, Hitati Works,
Ibarakiken, Japan.
March 1936.

III. *On the Atomic Forces of Solid State.*

By Weng Wen-Po *.

Part IV.

(12) In the last section of Part III.† we have assumed at the melting-point of crystals

$$\frac{df(r_0+x)}{dx}=0,$$

where $f(r_0+x)$ is the equivalent attractional force acting on an atom by a neighbouring one, and x is the displacement of the atom away from its own equilibrium position.

* Communicated by the Author. For Parts I. & II. see xxii. p. 49 (1936).

† Phil. Mag. xxii. p. 281 (1936).

It is probably due to the arbitrariness of the assumption or due to the internal complication of the atomic motion which makes the deviation in the value of $\left(\frac{br_0}{a}\right)$ by a factor 3/2. In general, we may assume that a crystalline solid begins to melt when the maximum displacement of the atom in any one direction reaches a certain fraction of the range at which the differential approaches zero. But we have shown that the atomic displacement at which $\frac{df(r_0+x)}{dx}=0$ is a constant fraction of the lattice spacing. It is equally good to assume that at the melting-point $x_m=\alpha r_0$, where x_m represents the maximum amplitude of oscillation in each direction, and α is a constant.

It should be emphasized that the displacement αr_0 is not necessarily the limit of displacement. It is, in general, possible for an atom in a crystal to travel farther than αr_0 without suffering collision, if atoms may collide.

Since, at the melting-point, the maximum amplitude in each direction is αr_0, the average kinetic energy in the direction will be $m\pi^2\nu^2\alpha^2 r_0$. Equating this to $\frac{1}{2}k_0 T_m$, where T_m is the melting-point in absolute scale, and k_0 is Boltzmann's constant, we obtain

$$\tfrac{1}{2}k_0 T_m = m\pi^2\nu^2\alpha^2 r_0. \quad . \quad . \quad . \quad . \quad . \quad (40)$$

Applying equations (9) and (10), as shown in Part I., *e.g.*,

$$\left.\begin{aligned} mn&=\rho, \\ n&=\rho N_0/W, \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad (9)$$

$$r_0 = kn^{-\frac{1}{3}}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

where m is the mass of an atom, n the number of atoms per cubic centimetre, ρ the density of the solid, N_0 Avogadro's number, and k a constant around 1·38, Lindemann's formula for the infra-red characteristic frequency will appear as

$$\nu = \sqrt{\frac{\overline{k}_0}{2}}\, \frac{N_0^{\frac{5}{6}}}{\alpha\pi k}\, T_m^{\frac{1}{2}} \rho^{\frac{1}{3}} W^{-\frac{5}{6}}. \quad . \quad . \quad . \quad . \quad (41)$$

From the result of Part III. we can easily find that $\alpha = \frac{3}{4}\left(\frac{a}{br_0}\right)$. If we let $\frac{br_0}{a}=17$, as a mean value for most

metals, the coefficient $\sqrt{\frac{k_0}{2}}\frac{N_0^{\frac{1}{3}}}{\alpha\pi k}$ will become $2{\cdot}8\times10^{12}$—the same value as given by Lindemann *.

When the maximum amplitude in each direction has just reached the range αr_0, the equilibrium of the atom under the binding force $f(r_0+\alpha r_0)$ will be unstable, although the magnitude of the restoring force $F(x)$ remains approximately in a straight line relation with x. Then, when the atom just travels to the point αr_0, it is in liquid state. But an atom having a displacement of αr_0 and zero kinetic energy will correspond to a liquid state at absolute zero temperature. Thus, owing to the interatomic conduction of energy, the atom cannot stay in liquid state, unless an amount of average kinetic energy corresponding to $\frac{1}{2}k_0T_m$ together with a mean displacement αr_0 is acquired. The amount of total energy required to fit this condition is the latent heat of fusion.

Assuming that the solid after fusion remains to be monoatomic, $f(r_0+x)$ and $F(x)$ can be still considered valid except that $F(x)$ is no longer a restoring force, but a force to pull the atom to any random direction. All the former conception of the stationary neighbouring atoms described in Part I. may be retained for their average effect, provided that the equilibrium point is changing from time to time. Some evidence of the linear force $F(x)$ in liquid state will be given in Part V. Since $F(x)$ is approximately linear to x, the average kinetic energy possessed by the atom at $x=\alpha r_0$ is equal to one-half of the total energy. Then the total energy required for the atom to remain in the liquid state and at the temperature T_m is $2m\pi^2\nu^2\alpha^2{r_0}^2$. Letting L be the latent heat of fusion per gram of substance, we have

$$Lm=2m\pi^2\nu^2\alpha^2{r_0}^2. \quad . \quad . \quad . \quad . \quad . \quad (42)$$

Eliminating ν from equations (40) and (42) we have

$$k_2T_m=mL.$$

Applying equation (9), we have

$$\frac{LW}{T_m}=k_0N.$$

* *Phys. Zeit.* ii. p. 609 (1910).

If we denote L_m, the latent heat of fusion, in heat unit, *i. e.*, calories per gram, the equation of latent heat will appear as

$$\frac{L_h W}{T_m} = R = 2{\cdot}0 \quad . \quad . \quad . \quad . \quad . \quad (43)$$

where R is the universal gas constant for one gram of substance in heat unit. The last equation represents a well-known empirical relation discovered by J. W. Richards. Equation (43) is similar to Dulong and Petit's law, which may be written as $S_h W = 3R$, where S_h is the specific heat at constant volume in heat unit. Both equations are either the direct or the indirect consequence of harmonic motion. Since the condition of harmonic motion is better fitted by solids at room-temperature than by liquids at melting-point, Dulong

TABLE XI.

Elements	$S_h W$*.	WL_h/T_m †.
Cu	5·92	1·99
Au	6·10	2·32
Na	7·06	1·7
Cd	6·26	2·53
Pb	6·43	1·96
Zn	6·14	2·60

and Petit's law is expected to hold better than equation (43). This is actually found to be the case, and a comparison of these two rules is illustrated in Table XI.

If a crystal does not melt into monatomic liquid, the atoms in solid will unite into molecules in fusion. Then the observed latent heat of fusion will be the difference between the theoretical value and the heat of molecular formation from lattice structure. Consequently the observed value must be less than the calculated value. A few examples are tabulated in Table XII.

Since fluorine ion is relatively small in size, it will be bound closer to the alkaline atom in forming molecules than chlorine. Thus the energy of molecular formation of NaF or KF will be larger than that of NaCl or KCl.

* *Handbuch der Experimental Physik*, xii. p. 193.
† *Handbuch der Experimental Physik*, xii. p. 592.

The numerical value so far as obtained in Table XII. is in good accordance with this consideration.

(13) From the same assumption that atoms are similar, $\left(\frac{br_0}{a}\right)$ will be nearly a constant for many kinds of sub-

TABLE XII.

Substance.	T_m °A.	L_h calc.	L_h obs.	Difference.
NaF	1265	1215	779	436
NaCl	1077	772	517	255
KF.....	1133	887	452	435
KCl	1045	558	310	248

stance, as has already been shown to be approximately true. In our previous work we have had

$$X_1^2 = \frac{m}{3a}\int_0^T S(T)\,dT, \quad \ldots \quad (15)$$

$$\int_0^T \alpha_l(T)\,dT = \frac{\Delta l}{l}, \quad \ldots \quad (25)$$

$$\frac{\Delta l}{l} = \frac{1}{2}\frac{b}{a}\frac{X_1^2}{r_0}, \quad \ldots \quad (22)$$

$$\nu = \frac{1}{2\pi}\sqrt{\frac{2a}{m}}, \quad \ldots \quad (8)$$

where X_1 is the maximum displacement in each of the three directions at a certain temperature T in absolute scale, $\frac{\Delta l}{l}$ the total elongation in length, and $\alpha_l(T)$ is the coefficient of linear thermal expansion at the temperature T. From the first three equations we can deduce

$$\frac{br_0}{a} = \left[\frac{\alpha_l(T)}{S(T)}\right]\frac{6kr_0}{n^{\frac{1}{3}}m}a. \quad \ldots \quad (44)$$

From equations (44) and (8), we get

$$\nu = \sqrt{\frac{\left(\frac{br_0}{a}\right)}{12\pi^2k^2}\left(\frac{N_0\rho}{W}\right)^{\frac{1}{3}}\left[\frac{S(T)}{\alpha_l(T)}\right]^{\frac{1}{3}}},$$

since $r_0 = kn^{-\frac{1}{3}}$ and $n = \frac{\rho N_0}{W}$. Taking $\frac{br_0}{a} = 17$, as before, we have

$$\nu = 2{\cdot}4 \times 10^7 W^{-\frac{1}{3}} \rho^{\frac{1}{3}} \left[\frac{S(T)}{\alpha_l(T)}\right]^{\frac{1}{3}}, \quad . \quad . \quad . \quad (45)$$

where the specific heat S(T) is in mechanical unit, namely, ergs per gram per degree. A very similar equation to equation (45) was obtained by Grüneisen in 1912, but the index of W in Grüneisen's formula is $l = -\frac{5}{6}$ *.

In Table XIII. the characteristic frequency of several metals calculated from equation (45) is tabulated. The value of $S(T)/\alpha_l$ (T) around 20° C. is adopted for the sake of convenience.

TABLE XIII.

Metals.	$\alpha_l(T)/S(T) \times 10^7$ gr. per joules.	$\nu . 10^{-12}$ calc.	$\nu . 10^{-12}$ accepted †.	Ratio.
Al	256	7·1	7·5	1·1
Co	293	7·5	8·2	1·1
Ni	292	7·5	8·2	1·1
Cu	431	6·0	6·7	1·1
Ag	809	3·9	4·4	1·1
Pt	655	4·5	4·4	1·0
Pb	2278	1·9	1·8	·9
Zn	863	3·9	4·4	1·1
Cd	1196	3·0	3·0	1·0

From Table XIII. the ratio of the accepted value to the calculated value shows that the coefficient of equation (45) is a little too small for most elements. The equation will fit better if the coefficient is replaced by $2{\cdot}6 \times 10^7$. If the specific heat is expressed in heat unit, we have

$$\nu = 1{\cdot}6 \times 10^{11} W^{-\frac{1}{3}} \rho^{\frac{1}{3}} \left[\frac{S_h(T)}{\alpha_l(T)}\right]^{\frac{1}{3}}, \quad . \quad . \quad . \quad (45')$$

where S (T) is the specific heat in calories per gram per degree at the temperature T. Equation (45′) can be also obtained by dimensional consideration ‡ with some suitable assumptions.

* *Ann. d. Phys.* xxxix. p. 293 (1912). *Cf.* Rawlins and Taylor, 'Infra-red Analysis of Molecular Structure,' p. 83 (1929).

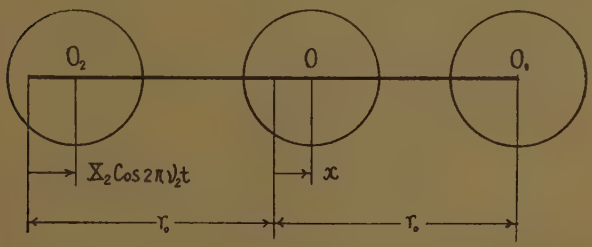
† Lewis, 'Physical Chemistry,' iii.

‡ In Grüneisen's original paper, it seems that atomic heat was used for specific heat in the numerical calculation, so his equation is equivalent to ours.

(14) In gases the molecules do not have equal velocity; the velocity is distributed according to certain law. In solids not only is the velocity of atoms unequal, but also the oscillating frequency cannot be unique for all atoms. The discrepancy of frequency between atoms can be regarded as the cause of certain property of solids.

The discrepancy from atom to atom may be due to the fact that the orbital motion of the electron is not synchronized. Thus a fluctuation of atomic force may result. This effect is small because of the comparatively large number of electrons in one atom. The outstanding factor which determines the frequency discrepancy is the b-term in the function $f(r_0+x)$ when the neighbouring atoms are also oscillating.

Fig. 3.



As the problem involved many particles a completely general solution is hard to carry out. So we will try to obtain an approximate solution to estimate the order of magnitude. Referring to fig. 3, let O_1, O_2, and O be three collinear adjacent atoms which are each r_0 apart from their neighbours respectively. As we have mentioned in Part I., these atoms may be treated as fictitious atoms so that the system represents the average case. Again, let O_1 be stationary, O_2 be in harmonic motion along the direction O_1, O_2 with an instantaneous amplitude $X_2 \cos 2\pi\nu_2 t$. Then the motion of the atom O will be defined by

$$m\frac{d^2x}{dt^2} = f(r_0 - X_2 \cos 2\pi\nu_2 t + x) - f(r_0 - x),$$

where x is the displacement of the atom O along O_1 O_2 from its equilibrium point, taking the direction toward O_1 as positive. $f(r_0 - X_2 \cos 2\pi\nu_2 t + x)$ represents the

coupling force between atoms O_1 and O, and $f(r_0-x)$ is that between O_2 and O. Here an allowable approximation has been made by considering $f(r)$ to be originated from a single neighbouring atom. In case of heteropolar binary compounds in cubic lattice, such as alkaline halides, the effect of other atoms besides the adjacent ones is around ten per cent. of the total only. Expanding $f(r_0+x)$ into $-ax+bx^2 \ldots$. we have

$$m\frac{d^2x}{dt^2} = -ax-bx^2-a(x-X_2\cos 2\pi\nu_2 t) + b(X_2\cos 2\pi\nu_2 t-x)^2,$$

or

$$m\frac{d^2x}{dt^2}+(2a+2bX_2\cos 2\pi\nu_2 t)x = aX_2\cos 2\pi\nu_2 t + bX_2^2\cos^2 2\pi\nu_2 t. \quad . \quad . \quad (46)$$

If we neglect $(2bX_2\cos 2\pi\nu_2 t)x$ in the left-hand member, the solution will be

$$x = c_1\sin(2\pi\nu t+c_2)+\frac{\nu^2X_2\cos 2\pi\nu_2 t}{2(\nu^2-\nu_2^2)} + \frac{b}{4a}X_2^2-\frac{b}{4a}\frac{\nu^2X_2^2\cos 4\pi\nu_2 t}{(4\nu_2^2-\nu^2)},$$

where $\nu=\frac{1}{2\pi}\sqrt{\frac{2a}{m}}$ and c_1 and c_2 are constants. The complementary solution determines a kind of harmonic motion which has already been discussed in detail. The particular integral $\frac{\nu^2X_2\cos 2\pi\nu_2 t}{2(\nu^2-\nu_2^2)}$ shows a coupling between the two atoms, and also gives the magnitude of the forced oscillation of the atom O by its neighbour O_2. Obviously this integral is the course of thermal conduction through atom. At low temperature, when the amplitude of atomic oscillation is generally small, the difference in frequency between two neighbouring atoms will be small and the thermal conductivity will be large. This is actually found to be the case for many non-metallic crystals. The particular integral $\frac{b}{4a}X_2^2$ is the amount by which the central point of oscillation is shifted away from the oscillating atom O_2, and the shift is evidently connected with the thermal expansion. As we have

assumed that atoms O_1 and O_2 are fixed, the shift of the atom O from atom O_2 due to the oscillation of atom O_2 is partially diminished by a repulsion of the atom O_1. In the actual case the position of O_1 is movable, we can find that the true thermal expansion will be twice this value, *i. e.*, $\frac{b}{2a} X_2^2$, the same value which we have found in Part II. The last particular integral $-\frac{b}{4a} \frac{\nu^2 X_2^2 \cos 4\pi\nu_2 t}{(4\nu_2^2-\nu^2)}$ defines overtones.

The thermal conductivity through atom, according to our explanation, depends largely on the discrepancy of the frequency in atoms, while the latter rises in the term $(+2bX_2 \cos 2\pi\nu_2 t)x$ in the left-hand side of equation (47). So the mechanism of atomic conduction of heat cannot be understood until we get the particular integral of the equation

$$m\frac{d^2x}{dt^2} + (2a + 2bX_2 \cos 2\pi\nu_2 t)x = aX_2 \cos 2\pi\nu_2 t, \quad (48)$$

which is a special case of equation (46).

The complementary integral of equation (48) is far less important than its particular integral. Suppose the atom O is connected to an infinitely large energy reservoir at absolute zero temperature which absorbs all the potential energy of the atom O from time to time, since the relation between atoms is a function of interatomic distance and independent of velocity, and O_2 to another at a temperature T corresponding to its amplitude of oscillation X_2, so that the conduction of energy is unidirectional. The rate of conduction of energy may be considered as the ratio of the total energy associated with the maximum amplitude of the forced oscillation to the time just sufficient for the atom O to acquire this amplitude from rest. After the atom O acquired its maximum amplitude its energy may be supposed to be totally absorbed by the energy reservoir, and the process repeats itself.

As the atom O_2 is coupled with many neighbouring atoms, the influence of coupling on the vibrational frequency neutralizes each other. Thus its frequency may be assigned as the mean frequency of the crystal,

i. e., $\nu_2=\frac{1}{2\pi}\sqrt{\frac{2a}{m}}$. Furthermore, let us define an auxiliary function $\phi(t)$ as

$$x=\phi(t)\cos 2\pi\nu_2 t. \quad . \quad . \quad . \quad . \quad . \quad (49)$$

The function $\phi(t)$ may be considered as the instantaneous amplitude of the forced vibration which is always positive when the temperature of the atom O_2 is not too high. Equation (48) is then transformed into

$$\frac{d^2\phi(t)}{dt^2}-4\pi\nu_2\tan 2\pi\nu_2 t\frac{d\phi(t)}{d(t)}+\frac{2bX_2}{m}\phi(t)\cos 2\pi\nu_2 t=\frac{aX_2}{m}. \quad . \quad . \quad . \quad (50)$$

$\phi(t)$ can be expanded into a special form of Fourier series, assuming it to be periodic and symmetrical, as

$$\phi(t)=a_0+a_1\sin k_1 t+a_2\sin k_2 t+\ \ldots\ +a_n\sin k_n t+\ \ldots$$

The coefficients can be evaluated by substituting the expansion into equation (50) thus,

$$\begin{aligned}&-(a_1{k_1}^2\sin k_1 t+a_2{k_2}^2\sin k_2 t\ \ldots\ \ldots\ +a_n{k_n}^2\sin k_n t+\ \ldots)\\&-4\pi\nu_2\tan 2\pi\nu_2 t(a_1k_1\cos k_1 t+a_2k_2\cos k_2 t+\ \ldots\\&\qquad\qquad +a_nk_n\cos k_n t+\ \ldots)\\&+\frac{2bX_2}{m}\cos 2\pi\nu_2 t(a_0+a_1\sin k_1 t+a_2\sin k_2 t+\ \ldots\\&\qquad\qquad +a_n\sin kt+\ \ldots)\\&=\frac{aX_2}{m},\end{aligned}$$

which is true for every value of t. When $t=0$, we get $a_0=\frac{a}{2b}$. Since only the average value of the amplitude $\phi(t)$ for a sufficiently long time is of interest, and the mean value of $\sin k_n t$ is zero, it is of no use to evaluate other constants. So we may conclude rigorously that the mean amplitude of forced oscillation is $\frac{a}{2b}$, which is independent of temperature.

The maximum amplitude of forced oscillation can be acquired in one-quarter of the period of the impressing oscillation, *i. e.*, $\frac{1}{4\nu}$ second. According to the foregoing explanation, in $1/4\nu$ second an amount of energy $2\pi^2 m\nu^2{a_0}^2$

is conducted through the atoms, and the rate of energy conduction per second per pair of atom will be $2\pi^2 m\nu^3\left(\frac{a}{b}\right)^2$ ergs per second.

As the energy is not a directed quantity, all immediate neighbours at one side with respect to a certain axis will contribute the same amount of conduction. For alkaline halides along one of the principal axes, there are four neighbours of the same mass and constitution as the atom, and five of different mass and constitution as shown in

Fig. 4.

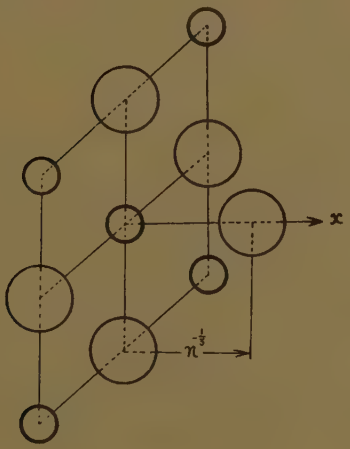


fig. 4. Moreover, the natural frequency of these two kinds of atoms is expected to be different. From the visibility curve of the reststrahlen, Rubens and Hollnagel * and others were able to identify the multiplet structure of the reststrahlen. The two components of the doublet in alkaline halides spectra are probably the two distinct frequencies of alkaline ions and halogen ions. Since the frequency difference between the two different ions is fairly large, the mutual conduction can be neglected. In other words the two vibrating systems are not in resonance. Thus the gain in energy of the atom

* *Verh. d. Phys. Ges.* xii. p. 83 (1910).

per second is only four times the value of $2\pi^2 m\nu^3\left(\frac{a}{b}\right)^2$, *i. e.*, $8\pi^2 m\nu^3\left(\frac{a}{b}\right)^2$. This quantity is equal to the product of the conductivity, the area, and the temperature gradient. Let the conductivity be K. The area in this case will be equal to $n^{-\frac{2}{3}}$ and the temperature gradient $T/n^{-\frac{1}{3}}$. Hence by equating these quantities, we have

$$K = 8\pi^2 m\nu^3\left(\frac{a}{br_0}\right)^2 r_0^2 n^{\frac{1}{3}}\frac{1}{T} \text{ erg/cm./sec.} \quad . \quad . \quad (51)$$

Assume $\frac{br_0}{a} = 14$ for alkaline halides as already estimated from the elastic data in Table VII. Applying equations (9) and (10), equation (51) can be written in the following form :

$$K = \frac{8\pi^2 k^2}{\left(\frac{br_0}{a}\right)^2 N_0^{\frac{4}{3}}} . W^{\frac{4}{3}}\rho^{-\frac{1}{3}}\nu^3 T^{-1}$$

$$= 1{\cdot}5 \times 10^{-32} W^{\frac{4}{3}}\rho^{-\frac{1}{3}}\nu^3 T^{-1} \text{ erg/cm./sec.}$$

$$= 1{\cdot}5 \times 10^{-39} W^{\frac{4}{3}}\rho^{-\frac{1}{3}}\nu^3 T^{-1} \text{ watt/cm./sec.} \quad . \quad . \quad (52)$$

Considering the conductivity of two kinds of atoms being additive, the value m appearing in equation (51) will be the average value of the two kinds, and in the same equation, r_0 and $n^{-\frac{1}{3}}$ can also be regarded as the average value. Thus in equation (52) W represents the average atomic weight of the constituent elements.

From equation (52) the conductivity is expected to be inversely proportional to the absolute temperature, which is a well-known fact and sometimes described as Eucken's law. Debye *, Schrödinger †, Compton ‡, and others have also proved this fact by various assumptions. Pauli §, however, pointed out that it is necessary to take into account the atomic nature of solid lattice in contradiction to Debye's and Compton's treatment.

In Table XIV. conductivity of several alkaline halides at zero degree centigrade is calculated by equation (52). Observed values are also tabulated in the last column.

* 'Wolfskehl-Vorträge zu Göttingen' (Leipzig, 1914).
† *Ann. d. Physik*, xliv. p. 916 (1914).
‡ Phys. Rev. vii. p. 341 (1916).
§ *Verh. d. Phys. Ges.* vi. p. 10 (1925).

The calculated values have the right order of magnitude. It is, perhaps, worthy of mention that equation (52) has no correction term due to quantization of energy even at very low temperature, because the equation contains neither specific heat nor any other energy function.

TABLE XIV.

Halides.	W.	ρ.	$\nu.10^{-12}$.	K(273° Å.) watt/cm. sec. calc.	K(273° Å.) watt/cm. sec. obs.*.
NaCl	29·2	2·16	5·8	·073	·070
KCl	37·3	2·65	4·8	·053	·070
KI	82·9	3·12	3·1	·041	·050

PART V.—*Extended to Liquid State.*

(15) When an elementary solid, such as metals, melts into monatomic liquid without chemical association between atoms, the interatomic force, namely $f(r)$, is still unchanged, although failing to hold the atoms in fixed relative position due to the large amount of kinetic energy possessed by them. Referring to a certain axis, the component of motion along the axis, owing to the presence of $f(r)$ still has the characteristics of harmonic oscillation. We may assume that after an atom travels a mean path along the axis it turns back. The path is not always the same so that the atom will drift gradually to one direction. In the path there is a point at which the component of $f(r)$ along the axis due to neighbouring atoms at two sides just neutralizes each other, and the maximum velocity results. At the extremity of the path the force $F(x)$, which may be defined as the algebraic sum of the forces $f(r)$ exerted by atoms at two sides along x-axis, will become sufficiently large to drive the atom back; then the velocity vanishes, and potential energy comes to a maximum. The atomic motion in these liquids will thus still be harmonic in character with the central point shifting continually, and the restoring force fluctuating widely.

The harmonic character of atomic motion in liquids can be shown by its physical properties. In the first place, equation (43), which is deduced under the assumption

* 'International Critical Tables,' v.

of harmonic motion, has been proved to be approximately true. In the second place, the product of the specific heat at constant volume for liquid metals and its atomic weight equals approximately 3R*, a direct conclusion of harmonic motion. And, in the third place, Haber's † square root formula for the infra-red characteristic frequency and that in ultra-violet is also true for liquids. Haber's formula can be deduced by assuming the molecules being in harmonic motion, the molecule and the electron vibrating independently. These evidences can scarcely be ascribed to chance.

(16) The presence of the force $f(r)$ in liquid becomes at once evident at the surface of the liquid. Perpendicular to the liquid surface the force keeps the atoms closely together, so that the density of the surface becomes larger than that of the bulk. Along the liquid surface the force causes a stretching on the surface giving rise to the phenomenon, surface tension.

Let us now treat the simple case of the surface tension of liquid a metal at its melting-point. In this case the trouble of taking the thermal expansion of the liquid into consideration can be avoided. First of all, it is necessary for us to consider the density and the thickness of the surface. The density of surface, as we have just mentioned, is greater than that of the bulk. But the density approaches a maximum value at which the distance between adjacent atoms is $(1-\alpha)r_0$, where α is a fraction already defined in Part IV. Over the maximum value the surface atoms would stick to each other so that the film becomes a material foil. Considering the comparatively great magnitude of the normal force which pulls the surface atoms together tightly, we may sensibly assume that the density on the surface reaches actually the maximum.

Various reasons indicate that the effective thickness of a surface film is not much more than the diameter of one molecular field, *i. e.*, the distance between the centres of two adjacent molecules. So, in a somewhat arbitrary manner, yet consistent with our former conception, we may assume that the surface tension can be estimated by a static model of atomic arrangement as shown in

* Eucken, 'Lehrbuch der Chemischen Physik,' p. 184.
† *Verh. d. Phys. Ges.* xiii. p. 1117 (1911).

fig. 5. This figure is schematically drawn to bring out the fact that the average distance between two adjacent atoms is r_0 in the bulk and $(1-\alpha)r_0$ on the surface.

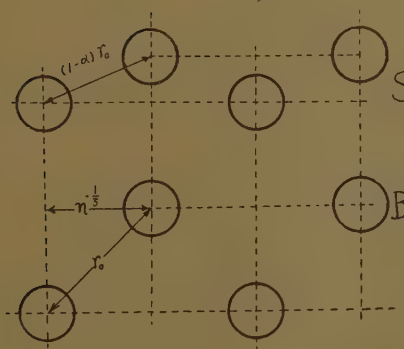
The force between two atoms on the surface equals approximately to $-f(r_0+\alpha r_0)$. Since in one centimetre length there are $n^{\frac{1}{3}}$ atoms, as shown in fig. 5, the surface tension will be equal to $-\frac{f(r_0+\alpha r_0).}{n^{-\frac{1}{3}}}$ We have

$$f(r_0+dr_0)=-ar_0\alpha\left[1-\left(\frac{br_0}{a}\right)\alpha\right],$$

thus

$$S_m=\frac{ar_0\alpha}{n^{-\frac{1}{3}}}\left[1-\left(\frac{br_0}{a}\right)\alpha\right]. \quad . \quad . \quad . \quad . \quad . \quad (53)$$

Fig. 5.



where S_m is the surface tension in dynes per centimetre at the melting-point. For metals, $\alpha=\frac{3}{4}\left(\frac{a}{br_0}\right)=\frac{3}{4\times 17}$ as obtained in Part IV., and $r_0=kn^{-\frac{1}{3}}$ where $k=1{\cdot}38$. Equation (53) will be $S_m=0{\cdot}015\,a$. We have proved in Part I. $\nu=\frac{1}{2\pi}\sqrt{\frac{2a}{m}}$, where ν is the characteristic infra-red frequency of the substance in solid state, and $m=W/N_0$, where W is the atomic weight and N_0 is Avogadro's number. Then we have

$$S_m=5{\cdot}0\times 10^{-25}\nu^2 W \text{ dynes/cm.}, \quad . \quad . \quad . \quad (54)$$

or

$$\nu^2=2{\cdot}0\times 10^{-24}\frac{S_m}{W}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (55)$$

Unfortunately, accurate determinations of the surface tension of metals are quite inadequate. Hence equation (55) can only be verified in a few cases. In Table XV. ν for several metals whose surface tension at their respective melting-point is accurately known was calculated by means of equation (55), and the results were compared with those from Lindemann's equation.

Owing to the more or less arbitrariness of our assumptions, the coefficient of equation (55) gives the values of ν always too large by 10 per cent. To allow for this factor, we may form a semi-empirical equation as follows:

$$\nu = 1{\cdot}27 \times 10^{12} S^{\frac{1}{3}} W^{-\frac{1}{3}}. \quad . \quad . \quad . \quad . \quad (55')$$

A complete verification of the equation still needs more experimental data.

TABLE XV.

Metal.	S_m.	10^{-12} equation (55).	10^{-12} Lindemann.	Ratio.
Pb	441*	2·0	1·8	1·1
Sn	500*	2·9	2·3	1·2
Hg	492†	2·2	2·0 ‖	1·1
Cd	628‡	3·3	3·0	1·1
Bi.....	376‡	1·9	1·7	1·1
Ga	358‡	3·2	2·6	1·2
Zn	753§	4·8	4·4	1·1

(17) We have seen that all kinds of atom are more or less similarly formed, so that their atomic force can be estimated by their lattice spacing. This principle may be extended to non-metals. Let us assume all nonmetallic elements are monatomic in constitution, near and under their melting-point. Applying all our foregoing treatment, the characteristic infra-red vibrational frequency will obey, at least approximately, the following equations:

$$\nu \propto W^{-\frac{1}{3}} \rho^{\frac{1}{3}} T_m^{\frac{1}{3}}.$$

$$\nu \propto S_m^{\frac{1}{3}} W^{-\frac{1}{3}}.$$

$$\nu \propto W^{-\frac{1}{3}} \rho^{-\frac{1}{3}} \beta_0^{-\frac{1}{3}}.$$

* L. L. Bircumshaw, Phil. Mag. xvii. p. 181 (1934).
† L. L. Bircumshaw, Phil. Mag. xii. p. 596 (1931).
‡ 'International Critical Tables,' i.
§ Thorfin R. Hogness, Journ. American Chem. Soc. xliii. p. 1621 (1921).
‖ Calculated from Debye's specific heat formula.

Since the surface tension and the melting-point of most non-metallic elements are known, we can find a verification in the first two equations. By eliminating ν, we get

$$\frac{\rho^{\frac{2}{3}}T_m}{W^{\frac{2}{3}}S_m} = \text{constant}$$

this quantity is listed in Table XVI. for several non-metals.

TABLE XVI.

Elements.	T_m°A.	$\rho^{\frac{2}{3}}$.	$W^{\frac{2}{3}}$.	S_m*.	$\frac{\rho^{\frac{2}{3}}T_m}{W^{\frac{2}{3}}S_m}$
A	84	1·391	11·7	12·5	0·80
Br	266	2·261	18·6	36	0·90
Cl	171	1·532	10·78	27	0·90
H	13·9	0·186	1·004	1·91	1·3
N	63	1·019	5·81	8·85	1·25
O	55	1·296	6·35	13·2	0·83
S	392	1·687	10·08	60	1·04
Se.....	493	2·793	18·4	72	1·04

It is very interesting to notice that the agreement is quite good, and the constants are all nearly unity. Further investigation on the subject will be recorded in future reports.

Imperial College of Science and Technology,
London.

IV. *Colorimetry with a Spectrometer.* *By* R. A. HOUSTOUN, *D.Sc., F.Inst.P., and* ALEXANDER J. YOUNGER, *M.A., B.Sc., Lecturers in the University of Glasgow.*†

THIS article is the fourth of a series. The first [1], by Houstoun, described a method of colorimetry in which the ordinary and extraordinary spectra produced by a glass prism and a quartz prism in series were superposed, and viewed through an eyepiece slit and nicol. The second [2], by Houstoun, described a method which

* 'International Critical Tables,' i.

† Communicated by the Authors.

was an improvement on the first, involving the reflexion on the second surface of an Iceland spar prism of a uniform white light which was seen superposed on one of the colours of the ordinary spectrum. The third [3], by Houstoun and Younger, was concerned with the investigation of discrepancies revealed in the second. These were found to be due to Houstoun's colour vision, and it was at the same time made clear that the small differences occurring in the colour vision of normal observers were of great importance for colorimetry and might lead to serious errors. The present paper is concerned with a modification of the method described in the second paper, which has been devised to minimize these errors.

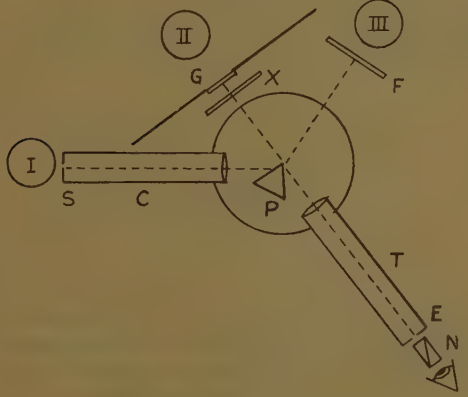
(1) The C.I.E. system of colorimetry is based on three primary colours and a white. It is necessary for all observers to agree at these four points. There are three methods of direct colorimetry—the vector method, the trichromatic method, and the monochromatic plus white method. In the second paper above referred to the vector method was employed. It is not easy on the basis of this method to obtain agreement of the observers at the white point. On the trichromatic method agreement can be obtained by a special calibration of the instrument for each observer; but the monochromatic plus white method gives agreement automatically not only at the four fixed points but at all points of the spectrum locus. We have therefore abandoned the vector method and adopted the monochromatic plus white method as a general procedure. The apparatus, which is shown in plan in fig. 1, is the same as before ; it is merely a question of using it differently.

P is an Iceland spar prism cut with its refracting edge parallel to the optic axis. A prism of this nature forms two spectra. Only the more deviated spectrum is used ; it is plane polarized in a vertical plane. E is a slit in front of which is rotated a nicol, N, with square ends. III is an opal Osram lamp. The eye looking in at E sees the second surface of the prism illuminated by one of the colours of the spectrum, and superposed on this by reflexion the white bulb of lamp III. The latter is polarized by reflexion in the horizontal plane. As the spectrum colour and white are plane polarized at right angles to each other their relative proportion in the mixture can

be altered by rotating the nicol. Thus the colour of the face of the prism can be changed from white to the spectrum colour in question.

The colour to be measured is in the case represented a plate of coloured glass X illuminated by lamp II through a piece of matt glass, and it is seen beyond the edge of the prism, the latter forming a sharp line of separation. Its brightness can be altered by moving lamp II. The operation of matching consists in moving the telescope until the hue is right, and then rotating the nicol until the saturation is right.

Fig. 1.



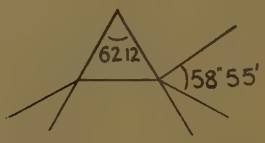
(2) In the first description of the apparatus it was mentioned that the dispersion of the spar prism was rather small, and it was suggested that it might be advantageous to use a glass prism and a spar prism in series. But the light from lamp III is then in some cases used at too great a distance from the polarizing angle to be extinguished by the nicol N. We have therefore adhered to the single spar prism. Two changes have, however, been made. Previously the refracting angle was 60°; it has now been changed to 62° 12′. Then, when the prism is at minimum deviation for the sodium yellow of the more deviated spectrum, the ray suffering minimum deviation leaves the prism at 58° 55′ (*cf.* fig. 2), the angle of polarization for sodium yellow, and a better

extinction of the reflected beam is obtained. Also the slow-motion screw of the telescope has been fitted with a micrometer drumhead, as the vernier on the circle was hardly accurate enough. The following table then gives the data for the calibration of the spectrum :—

Na	5893 Å.	49·70	He	5876 Å.	49·50
Tl.....	5351	43·55	He	5016	38·65
Li.....	6708	55·40	He	4713	33·27
He	7065	56·90	He	4471	27·97
He	6678	55·25			

(3) According to the C.I.E. scheme all observers must agree with the standard observer at illuminant B. This is a gas-filled lamp operated at a colour-temperature of 2848° K. used in combination with a filter consisting

Fig. 2.



of a layer, 1 centimetre thick, of each of two solutions contained in a double cell made of colourless optical glass, the Davis-Gibson solutions. The combination is designed to give a source of colour-temperature of about 4800° K, corresponding to one of the yellower phases of daylight. Solutions are, however, messy to work with, and one of the solutions contains pyridine, of whose keeping qualities we were doubtful. We attempted therefore to replace these solutions by something more convenient to handle, which would give approximately the same position on the colour diagram, and with this purpose in view Younger measured four samples of the Wratten No. 78 Tungsten to Daylight filter, an Osram daylight lamp, and various pieces of Chance's No. 9 glass-filter. Two pieces of the 78 filter had been in the department more than twelve years, of course carefully preserved in the dark ; the other two were bought together a month before they were tested. The four pieces of filter were found to vary so widely amongst themselves as to throw doubts on their

keeping qualities and put their use out of the question. The following are the other results :—

	x.	*z.*
Illuminant B. standard values	·348	·300
Daylight lamp	·438	·148
Chance's filter, 3 mm. piece	·398	·170
„ „ two 3 mm. pieces ..	·263	·365
„ „ two 4 mm. pieces ..	·280	·400

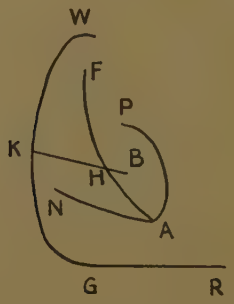
The daylight lamp is not blue enough. The pieces of glass 3 mm. thick were obviously from a different melting from the pieces 4 mm. thick. The measurements were made with the filters in front of a 220 volt gas-filled lamp operated at 250 volts. None of these substitutes proving suitable we fell back on the Davis-Gibson solutions. We find also that solutions made up more than a year ago are still giving the same point on the diagram.

(4) The monochromatic plus white method of colorimetry involves mixing a spectrum colour and a white, and it is necessary to know the relative brightness of the colours mixed. We first attempted to do this by direct heterochromatic photometry, by extinguishing lamp II and placing two screens on the prism table, one between the collimator and the prism and the other between lamp III and the prism, so as to cut off the lower half of the one beam and the upper half of the other. The eye looking through the slit then sees one-half of the prism face white and the other half coloured, and the intensities can be matched by turning the nicol. Results obtained in this way are of course not accurate ; it is difficult to compare the intensities of two surfaces when their colours are different. Following up an idea of Houstoun's, described in a letter to 'Nature' [4], we replaced the screens by pieces of fogged photographic film which transmitted the same fraction of the light incident on them. This diminishes the difference in colour of the two surfaces to be compared, and makes comparison easier ; but the method is not accurate when the difference is great, and cannot be regarded as a substitute for the flicker photometer. The flicker photometer was a complication

which we wished to avoid, and so, after spending some time on heterochromatic photometry, we abandoned it in favour of the following method.

The colour of lamp I was determined by Younger by the vector method. From its colour we obtained its colour-temperature, and its energy curve was calculated on the assumption that it was a black body. The ordinates were then multiplied by the visibility function and the slit width correction, and so the luminosity curve was obtained in arbitrary units. By means of the coordinate y the ordinates were then changed into their value for colour mixing purposes, also in arbitrary units. To bring them to the same units as were used for the white light

Fig. 3.



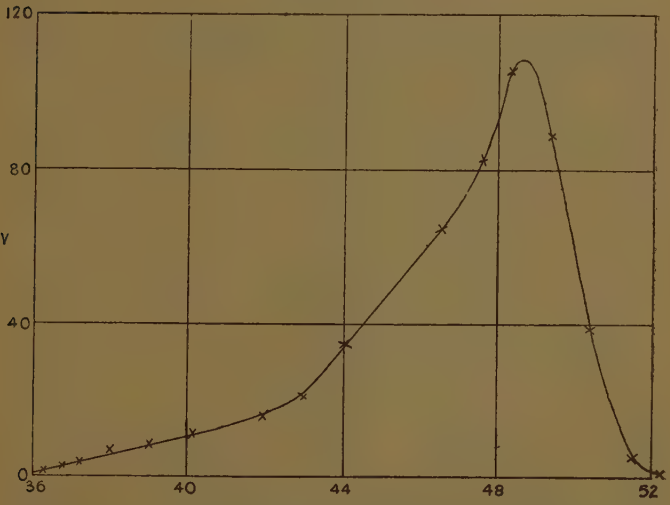
it was necessary to multiply by a correction factor ; this was determined by measuring a pale colour both by the vector method and the monochromatic plus white method.

This method, as has been stated, assumes that the energy curve of lamp I is that of a black body, and also that the light of this lamp suffers no selective absorption in its transmission through the instrument. The second assumption was tested by combining the light together again after its transmission and examining its colour, and seemed to be justified. This second method was, however, tested only superficially, being given up in favour of a third method which is undoubtedly much superior to its two predecessors.

This method is best explained by fig. 3. In this figure A and B represent the positions of two of the C.I.E. standard illuminants, the gas-filled lamp operated at 2848° K. and

the black body at 4800° K. RGW is the spectrum locus, the letters R, G, W denoting the positions of the colours red, green, and violet respectively. AN, AF, and AP are the nickel sulphate, copper sulphate, and potassium permanganate loci, *i. e.*, the positions of the colours that would be obtained by placing solutions of these salts of gradually increasing thickness in front of lamp A. To each point on a locus corresponds a definite thickness measured in centimetres.

Fig. 4.



The loci are first of all used to determine V, the value for colour mixing of each colour in the spectrum, the value for B being taken as unity. The thickness of copper sulphate which gives the point H, for example, is placed in front of lamp II, and telescope and nicol are adjusted to obtain a match, *i. e.*, the colours B and K are superposed to give H. Let α be the nicol reading. Then

$$\frac{HB}{KH} = V \tan^2 \alpha,$$

and since the ratio HB/KH is known V can be calculated. It is shown as a function of the telescope reading in fig. 4.

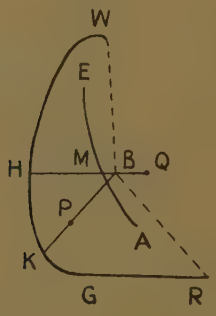
It varies with the slit width and also slightly with the position of lamp I before the slit, all the ordinates then being altered in the same ratio.

Let us suppose now that V is known and that it is required to measure an unknown colour. There are two cases. In the first case, which is represented by P (fig. 5), the colour lies within the region bounded by the dotted lines and the spectrum curve. When a match is obtained

$$\frac{PB}{KP} = V \tan^2 \alpha ;$$

V, α, and the position of K are known, and hence P is determined.

Fig. 5.



In the second case the colour is at Q outside the dotted lines. Here two settings are necessary. The unknown colour is placed in the position of lamp III and the standard illuminant B in place of lamp II, and the field matched. The equation is

$$\frac{BQ}{HB} = \frac{V}{V'} \tan^2 \beta,$$

where V′ is the value of the unknown colour. The point M, where HQ cuts one of the loci, *e. g.*, the copper sulphate locus, is obtained graphically, and the thickness corresponding to this point estimated. A solution of this thickness is placed in front of the illuminant A in the

position of lamp II, and a match again obtained. The equation is now

$$\frac{MQ}{HM} = \frac{V}{V'} \tan^2 \gamma.$$

On eliminating V'

$$\frac{BQ}{HB} \frac{HM}{MQ} = \frac{\tan^2 \beta}{\tan^2 \gamma},$$

and the ratio BQ/MQ is known; but the points B and M are known: hence Q.

In the settings described in this section it is not necessary that the intensities of both halves of the field should be the same, only that the colours should match.

(5) The following tables give the data from which the copper sulphate, nickel sulphate, and potassium permanganate loci were drawn :—

$CuSO_4 5H_2O$.—Strength 5·70 gm. per 100 c.c. solution.

Thickness.	x.	z.	λ.	σ.
cm.			mμ.	
1.....	·3781	·1982	568	·458
3.....	·2907	·2852	511	·315
6.....	·2201	·3815	497·5	·464
10.....	·1754	·4712	492·5	·551
15.....	·1492	·5462	489·5	·616

$NiSO_4 7H_2O$.—Strength 15·124 gm. per 100 c.c. solution.

Thickness.	x.	z.	λ.	σ.
cm.			mμ.	
1.....	·3576	·1703	560	·577
2.....	·3047	·1789	542	·605
6.....	·2069	·1926	520	·716
14.....	·1321	·2462	511	·807

K_2MnO_7.—Strength 0·0417 gm. per 100 c.c. solution.

Thickness.	x.	z.	λ.	σ.
cm.			mμ.	
½	·469	·179	607	+·307
1	·471	·189	616	+·348
2	·466	·219	497	−·529
3	·449	·267	502	−·638
4	·420	·340	516	−·800
6	·376	·382	544	−·861
10	·302	·436	440	+·011

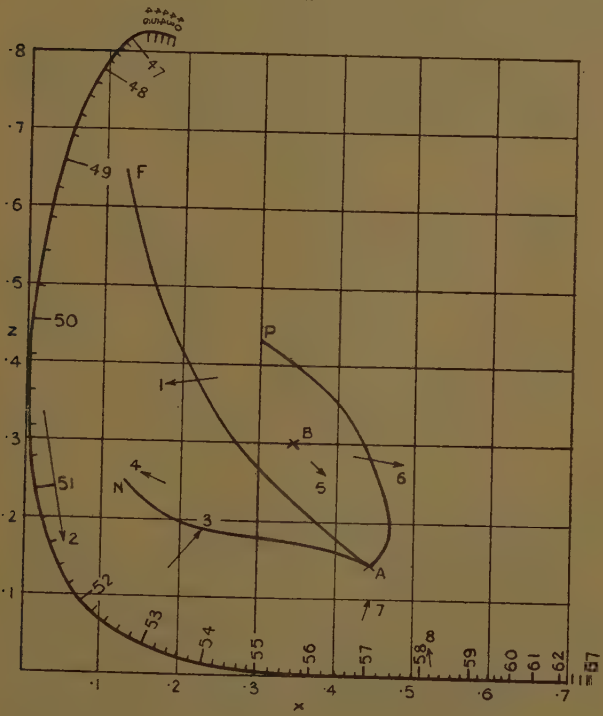
The values for copper sulphate were calculated from the transmission coefficients given by K. S. Gibson [5]. The values for nickel sulphate were calculated from absorption coefficients obtained by Houstoun and Russell [6] and Houstoun [7] more than twenty-five years ago. The values for potassium permanganate were determined by Younger by the vector method. The tables are put forward meantime provisionally; more points should be calculated for copper sulphate, and the absorption coefficients for nickel sulphate and potassium permanganate should be determined with modern accuracy. It is also not certain that potassium permanganate solutions are sufficiently permanent. The values for hue and saturation were obtained graphically, and are consequently only approximate.

Thicknesses of different solutions may be combined. For example, 6 cm. $CuSO_45H_2O$ and 10 cm. K_2MnO_7 in series in the above-mentioned strengths give the point $x=·268$, $z=·416$.

(6) After the new method had been worked out and was in order observations were made on eight test colours by Houstoun and Younger in order to see by how much they would differ. The results are shown by the arrows in fig. 6, the head of the arrow denoting Houstoun's reading and the tail Younger's. Each observer made two readings for each point, and the readings were taken alternately. The copper sulphate, nickel sulphate, and potassium permanganate loci are also entered on the

diagram. The want of agreement is very depressing when we consider that the method brings the two observers automatically into agreement not only at white and the three primaries but at every point of the spectrum locus, and that the observations were made under identical conditions, so that the differences are due to nothing

Fig. 6.



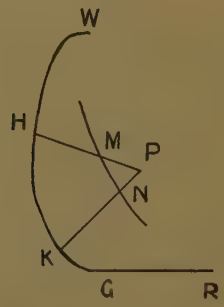
else but the observer's colour vision. We are afraid that such differences are by no means uncommon.

There are, however, two mitigating circumstances to be considered. The observations were reduced with Younger's value of V. If Houstoun had determined a value of V for himself the radial difference would have been much reduced. Also Younger did not always

determine V by the most suitable locus. For example, in the case of the magenta point denoted by σ the copper sulphate locus was used to determine V. It would be better to use the permanganate locus which passes much nearer the point.

This brings us to what we consider the solution of the problem, namely, to fix as many of these loci as possible by the indirect method and to determine V by means of the locus which passes nearest the point. The loci might conceivably also be used to determine wave-length, but it would lead too far to go into this matter at present. If V is determined in this way the observer is necessarily in agreement with the C.I.E. observer in the region under

Fig. 7.



consideration, and the method really becomes a combination of the direct and indirect methods.

As an example of the way in which the loci can be made to eliminate differences in colour vision fig. 7 may be considered. It represents the determination of a white by the vector method. All whites lie close to the copper sulphate locus. Two points M and N are taken on this locus near P, and each is combined with a spectrum colour to give P—that is, the lines HM and KN are fixed. The distances MP and NP are so short, that if the positions of H and K are slightly inaccurate, the position of P is not appreciably affected.

(7) The following table gives the values we have used for the Planckian locus. The x, z were calculated from a table given in other units in a paper by Deane B. Judd [8]. We give them here, as we have not seen them in print

anywhere, although they seem to be well known. The hue wave-length and saturation were obtained graphically.

Colour-temperature.	x.	z.	λ.	σ.
			mμ.	
100° K. .	·7348	·0	700	1·00
300	·7341	·0	686	1·00
600	·7091	·0	631	1·00
1000	·6519	·0030	609	·990
1500	·5852	·0212	594	·947
1900	·5371	·0514	590	·852
2360	·4893	·0958	586	·735
2848	·4474	·1448	584	·583
3500	·4051	·2046	582	·384
4800	·3508	·2932	571	·038
6500	·3132	·3634	484	·079
10000	·2805	·4312	480	·123
24000	·2531	·4937	478	·171
∞	·2400	·5260	476	·181

The Planckian locus is plotted in fig. 8, together with the copper sulphate, nickel sulphate, and potassium permanganate loci on a hue-saturation diagram. The Planckian locus is given by the dotted line. This diagram suffers from the disadvantage that equal displacements do not represent equal changes of hue, the discrepancy being very marked indeed; but it is easily understood, and the scale-numbers convey a meaning at once to the student of physics, whereas the x–z diagram is meaningless until the spectrum locus is laid down on it, and requires hours to understand. We are therefore experimenting with this method of representation. Incidentally it lends itself very well to the monochromatic plus white method, a horizontal displacement representing a movement of the telescope and a radial displacement a rotation of the nicol.

(8) Fig. 9 gives random measurements made of the colour of daylight at Glasgow during August on eight different days between the hours of 10 A.M. and 4 P.M

Fig. 8.

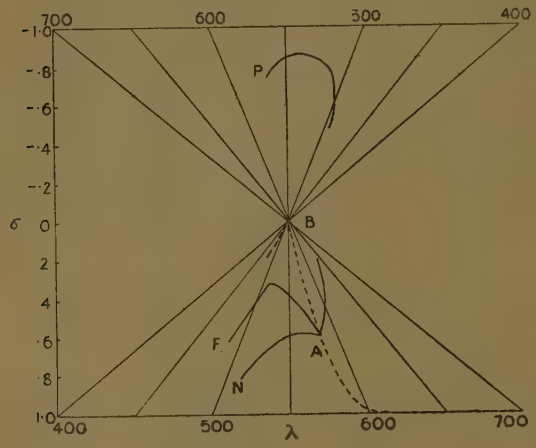
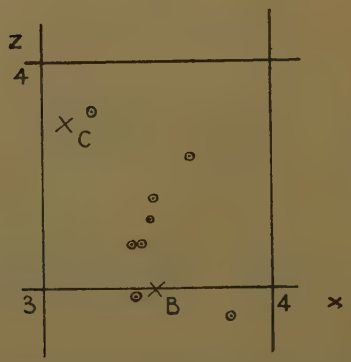


Fig. 9.



They give an idea of the variation to which it is liable, and show its unsuitability for colorimetric measurements.

The observations were made in the first instance with a Lovibond tintometer, allowing the daylight to fall on the magnesium carbonate surface supplied with the latter.

Then the glasses and lamp and blue glass were removed from the tintometer and set up in place of Lamp III in our arrangement. Copper sulphate solutions of two different strengths were next set up in succession in front of Lamp II, and the positions of the colours obtained by the vector method as indicated in fig. 7. As has already been stated this is a very trustworthy method of measuring a white.

The positions of the standard illuminants B and C are indicated by the two crosses, so our values lie roughly between them.

References.

(1) "A New Method of Colorimetry," Phil. Mag. xvi. p. 945 (1933).
(2) "Colorimetry with a Spectrometer," Phil. Mag. xvii. p. 1047 (1934).
(3) "The Personal Error in Colorimetry," Phil. Mag. xix. p. 1107 (1935).
(4) 'Nature,' cxxxv. p. 1000 (1935).
(5) Journ. Opt. Soc. America and Rev. of Sc. Inst. ix. p. 112 (1924).
(6) Proc. Roy. Soc. Edin. xxix. 2, p. 68 (1908).
(7) Proc. Roy. Soc. Edin. xxxi. 4, p. 538 (1910–11).
(8) Research Paper R.P. 756, 'Bureau of Standards.'

V. *Effect of the Age of Phenolite on the Coefficient of Photoelastic Extinction.* *By* J. Kuno, *Dr. Eng., Assistant Professor of Civil Engineering, Kyushu Imperial University, Fukuoka* *.

A SERIES of tests were carried out during the last twelve months in order to settle the question whether the photoelastic phenomenon in phenolite is affected by the age of the material or not. Six pieces of phenolite specimen were supplied by the Riken (Institute of Physical and Chemical Research), Tokyo. Each specimen is 20 mm. wide, 150 mm. long, and 6 mm. thick (see fig. 1 on p. 459 of Phil. Mag. xix., 1935). Testing was made at regular intervals, namely, on 20th of each month. The same specimen was examined only twice (2 months).

The specimen was held in the field of circularly polarized monochromatic light (5461 Å.), and was tested as a beam subjected to two equal loads at both ends of the middle-third of the span, the span-length being 120 mm. The

* Communicated by the Author.

total load P was applied to the specimen by means of a lever. The isochromatic lines at the centre of the span were photographed on six stages of the load for each specimen.

Stages....	1.	2.	3.	4.	5.	6.
P, kg.....	5	7·5	10	12·5	15	17·5

On each stage three operations were carried out, namely :—

1. To apply the load without vibration.
2. To photograph at one minute after loading.
3. To remove the load at once.

Repetition of such a procedure gave a system of data (which is similar to fig. 5 on pl. xvi. of Phil. Mag. xix., 1935).

The coefficient of photoelastic extinction, K_1 (creep of one minute being taken into account), was then calculated for each specimen by the method previously described (Phil. Mag. xix. p. 464, 1935). The results are as follows:—

Date of test.	Value of K_1.	Specimen.	Room-temperature.
Dec. 5	12·30	A	24° C.
Dec. 20	12·45	A	23
Jan. 21	12·85	B	16
Feb. 20	12·98	B	20
Mar. 20	13·10	C	18
Apr. 20	13·35	C	15
May 20	12·68	D	21
June 20	12·10	D	23
July 20	11·50	E	26
Aug. 20	11·85	E	29
Sept. 20	12·23	F	22
Oct. 20	12·22	F	18

From December to April the value of K_1 increases gradually, while it decreases after April. It reaches a minimum in July. Considered as a whole the variation of K_1 is comparatively small. It may therefore be concluded that the photoelastic property in phenolite is practically little affected by the age of the material so far as these experiments are concerned.

Taking the arithmetic mean of the above values we get for the phenolite tested

$$K_1 = 12{\cdot}47 \pm 0{\cdot}11 \text{ kg./cm.}$$

VI. *Rotation of the Earth and Foucault's Pendulums of Short Lengths.* *By* B. DASANNACHARYA, *M.A., Ph.D., F.Inst.P., Professor of Physics, Benares Hindu University, and* D. HEJMADI, *M.Sc.(Benares), Benares Hindu University, India* *.

IN 1851 Foucault † showed that the plane of oscillation of a simple pendulum continuously rotates by an amount θ depending on ω, the angular velocity of rotation of the earth about its own axis, and λ, the geographical latitude of the place of experimentation,

$$\theta = -\omega t \sin \lambda,$$

where t is the time during which the rotation is observed. This method has been the subject of several investigations. Firstly, to demonstrate the effect to the public, pendulums of length as great as 100 metres have been used with heavy bobs weighing as much as 28 kilogrammes. Here we get the advantage of great amplitude combined with a long period for the oscillation and duration for observation. The disadvantage arises from having to find tall towers and rigid supports. Attempts at diminishing the length have been made. We may mention the experiments of T. G. Bunt ‡ and Phillips. Kammerlingh Onnes used a physical pendulum 1·2 metres long, supported from a carefully constructed Cardani suspension, and swinging in vacuum. A. C. Longdon reported experiments in which copper-wire was used hung from double roller suspensions in bifilar arrangement. Their results will be discussed later.

From the point of view of theory opinion is divided as to how far Corioli's force due to the rotation of the earth is alone responsible for the magnitude and nature of the effects observed §.

* Communicated by the Authors.

† L. Foucault, *Compt. Rend.* xxxii. p. 135 (1851) ; Pogg. *Ann.* lxxxii. p. 458 (1851). For literature : *Handbuch d. Experimental Physik*, A. Haas, ii. p. 123 (Wien Harms Series). *Handb. d. Physik.* v. pp. 338 & 452 (1927) (Geiger u. Scheel Series), *Handb. d. Physik. u. Tech. Mechanik.* Auerbach and Hort, ii. p. 665 (1930).

‡ *Vide* p. 79, footnote.

§ Denizot, *Ann. d. Physik.* xviii. p. 299 (1905) ; xix. pp. 613, 868 (1906) ; *Phys. Zeit.* p. 507 (1906) ; vi. p. 667 (1905). M. Koppe, *Phys. Zeit,* pp. 604, 605 (1906). Tesar, *Phys. Zeit.* vii. pp. 199–207 (1906) ; vi. p. 556 (1905).

Previous investigators used mostly metallic wires for the pendulum. This necessitated the use of long wires and heavy bobs in order to minimize the effect of the rigidity of the suspension wire in causing parasitic oscillations. We have found that the snapping of wires and difficulty of rigid supports and finding suitable place can all be overcome by using torsionless strings of silk, or even cotton, of about a millimetre in diameter. These are easily procured in any market, as such strings are used in angling rods and for other sundry purposes. These strings have practically no rigidity, and hence small bobs and small lengths of strings could be used, and troubles due to yielding of the support consequently disappear.

In the experiments reported below we have been able to reduce the length of the pendulum to 95 cm. and work with bobs only 150 grms. in weight. No recourse to working in vacuum was found necessary. Usually it is found difficult to get a correct fixing of the string, and we have investigated in great detail the cause for this. Finally, we shall show how the mathematical theory could be extended to take into consideration the effect of air resistance on the growth of the elliptical orbit of the bob of the pendulum, and discuss the nature of the remaining corrections, which have not been overcome.

We have succeeded in so reducing the errors of support and rigidity of suspension that for the first time the pendulum bob is shown to go as required by theory in an orbit which is anticlockwise whatever the azimuth in which the pendulum may be set swinging, the Foucault angle remaining clockwise. This point will be fully discussed.

Experimental Part.

Determination of the General Nature of the Errors.

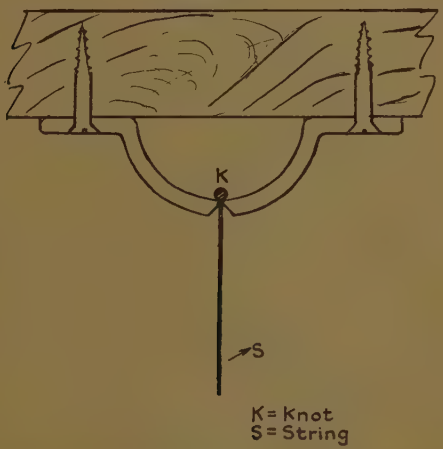
Apparatus.—On the top of a wooden table, 50×50 cm. and about a metre high, is built a wooden scaffolding about 275 cm. high. The top plank of wood, about 20×20 cm., serves the purpose of fixing the pendulum support. Weights were used to stabilize the table.

A woven silk-string, 245 cm. long, about 1 mm. diameter, was used. It was hung (see fig. 1) by passing it through a hole 1 mm. diameter in an iron piece, and prevented from slipping by a small knot K at that end of the string.

The weight of the bob was 450 gm., it is of steel and pear-shaped. A pin is attached to its pointed lower end in a hole and fixed with beeswax. The bob was attached to the string with a screw and hole.

The pendulum was loaded with an extra weight of about 500 gm. for about three hours to stretch it well when the string is being used for the first time only. The pendulum is started in the usual way by pulling the bob aside, supporting it about its centre of gravity with a loop of thread. When the bob is steady the thread is burnt outside of the loop and the oscillations start.

Fig. 1.



Knot support that acquires a favoured direction of vibration.

The movement of the shadow of the pin-point due to an electric bulb at the top at a little inclination (to the axis of suspension) is noted on a slip of millimetre graph paper. This paper could be rotated about a pin-point under the equilibrium position of the shadow-tip as the direction of rotation of the pendulum rotates. The amount of rotation is read off on a circular scale marked on paper and graduated in degrees.

At first on starting, the shadow moves along a straight line, but in a few minutes one could note that the orbit becomes an ellipse; the length of the minor axis can be

TABLE I.

Pendulum with favoured Direction of Vibration obtained by the use of Knot Suspensions.

Knot.	No.	Date 1935.	2P (cm.).	t (min.).	2 Q (cm.)	θ_Q.	t_γ (min.).	γ.	θ_γ.	t_r.	θ_r.
1.....	I (1)	15/2	54	40	4·1	24°	75	·5	58°	145	84°
	(2)	18/2	54	30	5·0	25	60	·42	55	120	82
	(3)	25/2	56	42	4·8	32	65	·52	55	125	90
	I′ (1)	21/2	68	25	3·5	−23	40	−·23	−39	85	−55
	(2)	22/2	56	25	3·5	−23	45	−·28	−40	90	−60
	(3)	27/2	49	20	3·7	−18	45	−·23	−40	85	−56
	I″	28/2	56	..	..	− 7*	..	..	..	..	..
	II (1)	11/3	58	30	4·9	28	60	·46	64	120	87
	(2)	11/3	58	30	4·9	26	60	·50	58	120	87
	(3)	13/3	58	30	4·9	24	60	·54	60	120	87
	II′ (1)	5/3	51	20	4·3	−22	37	−·21	−37	75	−56
	(2)	9/3	51	20	3·3	−22	30	−·22	−28	80	−55
	(3)	9/3	51	20	3·3	−22	30	−·22	−31	75	−54

2.....	III	(1)	17/3	56	30	2·4	15	60	·2	27	165	38
		(2)	18/3	56	30	2·6	15	50	·2	24	165	43
	III′	(1)	16/3	40	40	3·2	−26	87	−·45	−62	180	−102
		(2)	20/3	46	45	3·3	−32	80	−·42	−61	180	−102
3.....	IV		21/3	54	30	3·4	−29	55	−·29	−58	120	−86
	IV′		22/3	54	37	2·9	20	75	·36	38	196	62
4.....	V		22/3	56	23	5·0	32	37	·34	53	90	74
	V′		23/3	56	20	3·8	−28	35	−·23	−43	75	−67
	VI		28/3	56	23	1·1	−14	40	−·08	−20	75	−30
	VI′		28·3	56	60	3·2	22	100	66	83	210	120
	VII		29·3	56	15	1·6	6·5	30	·08	13	80	16
	VII′		29·3	56	60	4·4	−40	75	−·77	−74	150	−137

In I and I′ the initial directions of vibration are in directions perpendicular to each other. The same for II and II′, etc. I″ is for an intermediate position which makes 25° with that of I. In I″ the rotation is linear and is 7° per hour, the correct Foucault angle with correct sign. At time t_Q 2Q is maximum and the rotation is θ_Q; at t_γ the ratio of the minor to the major axis, namely γ, is maximum, at time t_r the maximum deflexion θ_r is attained, and γ passes through zero and reverses in sign.

* This is nearly the correct value of the F.A.

read off on the millimetre paper. The minor axis gradually increases, but its growth depends on many factors. The major axis, on the other hand, continuously diminishes, due to damping caused by air, and its rate is practically independent of the size of the minor axis. The ratio γ of the minor to the major axis is an important quantity, and its value is obtained by noting the lengths of the two axes at different intervals of time, usually 15 minutes. The direction of motion in the ellipse is very characteristic. Looked at from the point of suspension of the string at the top it may be clockwise (−) or anti-clockwise (+). The rotation θ of the major axis, like γ, is highly susceptible to errors in suspension. It could also be + or −. If all errors are eliminated as we shall show later the sign of γ is +, and according to some calculations should be hardly observable and should be independent of time. The value of θ is 6·4° per hour for the latitude of Benares, namely 25° 18′ N., and of − sign, clockwise. Both the quantities θ and γ should not depend on the initial plane of vibration.

In Table I. experiments with different knots are given.

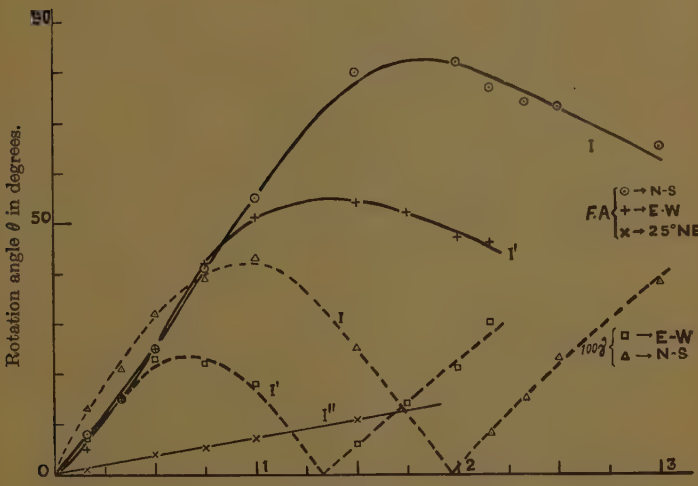
In I (1) the values for the rotation of the major axis are tabulated for three different conditions: firstly θ_Q when the minor axis 2 Q attains its maximum value, secondly θ_γ when the ratio γ of the minor to the major axis attains the maximum value, and finally $\theta_\gamma = \theta_m$ when 2 Q reverses its sign and θ has its maximum value. The times at which these different values are attained are also shown. Fig. 2 shows the curves I and I′ for two typical directions normal to each other. Coming back to Table I., I (1) and (2) have the same initial amplitude doubled, namely, 2P=54 cm. In (1) 2Q becomes 4·1 cm. in 40 minutes. At this time θ is 24° in an anti-clockwise direction. In 75 minutes from start γ attains maximum value +·5, the meaning of + is that the rotation in the ellipse is anti-clockwise; the value of θ_γ is 58° and is nearly twice θ_Q. The maximum value θ_r takes place 145 minutes after start, and its value is 84°. At this point minor axis is zero and changes sign. After this θ diminishes slowly. I (2) and (3) show that the values recorded in I (1) are reproducible. The maximum value for γ is about +·47, the direction of initial vibration is N.–S.

I′ (1), (2), and 3 are experiments repeated in a direction E.–W. The initial rotation in the orbit is now negative

and θ is also negative. The values of θ_Q, θ_γ, and θ_r are different, but the general course is the same. We now start the pendulum in a plane making an angle of 25° N.E. The values for θ both in magnitude and sign are the correct ones to be expected for a correctly adjusted Foucault pendulum. Even the vibration in the small ellipticity developed is of correct sign, namely, positive. The result is shown in I″ in Table I. and fig. 2.

In II we have got the repetition of I, however, with direction E.–W. and in II′ that of I′ consequently N.–S.,

Fig. 2.



Free support with favoured direction of vibration, rotation angles and γ.

obtained with the whole table bodily rotated by 90°. The values of θ show that the rotation of the table has not made any material change.

In III we again have for directions N.–S., using a fresh knot, knot no. 2. The values of θ are quite different from those in I. For example, θ_r now is nearly $+40°$ instead of $+85°$ obtained in I. In the normal direction E.–W. shown in III′ the values again are different; they are shown to be reproducible.

In IV and IV′ we have results for still another knot bigger than knot 2. In IV we now have for θ negative values, the directions being N.–S., but for the normal direction positive values, and those values are different again from those with previous knots.

With a very small knot, no. 4, results are obtained in perpendicular directions in V and V′ N.–S. and E.–W. respectively. In VI and VI′ these are given for perpendicular directions with the same knot, but its position in the top support is slightly rotated from its previous position in V and V′. In VII and VII′ we have for still another position of the knot in its suspension. These results with knot 4 show that the values obtained depend primarily on the position of the knot, on the suspension hole, and on the directions of vibration with respect to the knot, but *not on the directions in space.*

The general conclusions for these pendulums with suspension from knots are given below.

1. The sign of vibration (referred to as vibration) in the ellipse and that of the rotation of the major axis are the same for any direction of initial oscillation.

2. This sign changes to its opposite in the neighbouring quadrants.

3. There is an intermediate direction for which the minor axis of the ellipse can be very small, and nearly the correct value for θ as required by the theory of the Foucault's pendulum could be obtained (see I″ of 28.2.35).

4. The value of θ when the minor axis is max. (θ_Q), when the ratio of minor to major axis is max. (θ_γ), and finally when minor axis is zero, and the respective times at which these values are attained show (as *averaged* for *all* the trials shown under different knots, see Table II.) that

$$\frac{\theta_\gamma}{\theta_Q}=1{\cdot}99, \quad \frac{\theta_r}{\theta_\gamma}=1{\cdot}53, \quad \frac{\theta_r}{\theta_Q}=2{\cdot}99,$$

and the corresponding values for the times are

$$\frac{t_\gamma}{t_Q}=1{\cdot}80, \quad \frac{t_r}{t_\gamma}=2{\cdot}29, \quad \frac{t_r}{t_Q}=4{\cdot}06.$$

So a typical relation $\theta=f(t)$ would appear as follows.

At point 1 minor axis is max. at 2 at nearly double the time, θ also becoming double γ is max. and at point 3 at 4 times the time, minor axis reverses with θ becoming 3 times

its value at point 1. It also follows that up to 2 the curve is straight. The values of θ are in the ratio 1, 2, and 3 for times 1, 2, and 4.

5. The maximum value of θ is attained for a time at which the minor axis becomes zero. As the minor axis grows again with opposed sense of rotation the value of θ diminishes.

6. The numerical values of $\Sigma\theta_r = \theta_r + \theta_r'$ for two perpendicular directions are very nearly independent of the nature and position of the knot, and equal to 146° (Table II., column 4).

TABLE II.

Relations between the different Quantities shown in Table I.

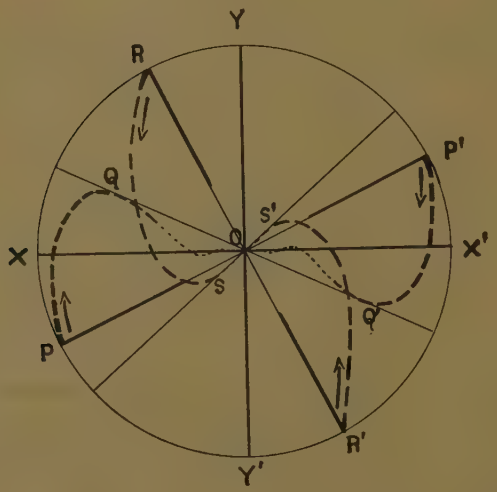
Knot.	$2\bar{\theta}_Q$.	$2\bar{\theta}_\gamma$.	$2\bar{\theta}_r$.	$\bar{\theta}_r/\bar{\theta}_Q$.	$\bar{\theta}_r/\bar{\theta}_\gamma$.	$\theta_r/\bar{\theta}_Q$.	$\bar{t}_\gamma/\bar{t}_Q$.	$\bar{t}_r/\bar{t}_\gamma$.	$\bar{t}_r/\bar{t}_Q$.	$2\bar{\gamma}$.
1	53°	94°	142°	1·94	1·52	2·92	1·81	2·09	3·79	·72
2	44	87	147	1·92	1·62	3·07	1·9	2·57	4·67	·64
3	49	96	148	1·95	1·5	3·0	1·9	2·35	4·4	·65
4	47	95	148	2·05	1·5	3·0	1·6	2·15	3·4	·72
	48·2	93	146	1·99	1·53	2·99	1·80	2·29	4·06	·68

7. The direction at which γ is max. defines a unique direction with respect to the knot, and can be called *a favoured direction of vibration for the knot.* Let X O X′ (fig. 3) be the most favoured direction brought about by some fault or dissymmetry in the knot suspension. Let O be the point which gives the position of the bob under free suspension. If the bob is set vibrating along P O P′, making an angle $\theta = \theta_\gamma$, then the plane will tend to align itself by the shortest route along X O X′, as that is the most favoured direction. The force with which it will be rotated into that direction is proportional to $\sin\theta$, and the rotation as well as the sign of vibration in the ellipse will be, as can be expected from the principle of the spherical pendulum, clockwise. The rotation will continue, however, beyond X O X′, due to the kinetic energy of motion, and will grow up to Q O Q′ till the restoring force of the favoured direction overcomes it. The ellipticity on the return journey will change its sign.

On account of damping the major axis diminishes continuously, and the intensity of the aligning force would consequently also decrease ; hence the transit back to XOX′ would be at a slow rate.

It is quite obvious that for an initial direction perpendicular to POP′ the alignment to the direction XOX′ would be anticlockwise at first, till it comes to rest on the opposite side of XOX′. Then there will be clockwise rotation for journey back to the favoured direction.

Fig. 3.



8. For any direction on either side of the favoured direction the sense of rotations are opposite.

9. In a direction normal to the favoured direction correct Foucault rotation can be obtained provided the experiment is not continued till θ becomes more than about 8°. This point has been settled after careful trials, not reported in the table.

10. If continued too long in the normal to the favoured direction the value of θ gets abnormally high, since the direction of the major axis in tending to align itself to the favoured direction has to fall very nearly 90°, and, further, the kinetic energy associated with the fall carries

the major axis to about 53 per cent. of the value of full, *i. e.*, 48° on the other side of the favoured direction. The total value of θ_r will therefore be $90°+48°=138°$. This is the basis for the value of 146° given under 6. This is a very important criterion. For example, if θ_r in any experiment is 30° then its value for a perpendicular direction of initial vibration should be $146-30=116$.

11. It would follow from 7 and 8 that $2\overline{\theta}_\gamma=\theta_\gamma+\theta_\gamma'$ would be very nearly 90° (column 3, Table II.).

12. If θ as measured from the favoured direction is very small the effect on rotation due to the most favoured direction would be negligible for short times initially (see 7), and we should get nearly correct Foucault angle and correct sign for the vibration, as has been carefully verified.

13. The numerical values of γ summed up for two perpendicular directions are also independent of the size of the knot and direction of vibration. This sum is very nearly 0·7 (last column, Table II.).

Application.

Experiments with rigidly fixed Supports.

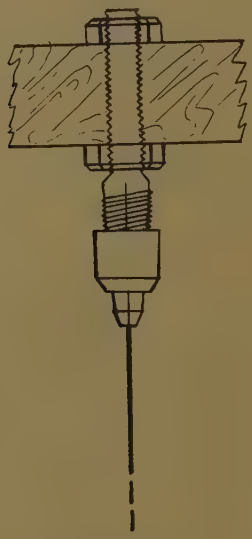
We want to know now to what extent the conclusions given above can be applied to the case when the knot is removed or it is well fixed with sealing-wax or with metal jaws.

It is found that nearly all the abnormalities disappear, and correct Foucault angle can be obtained with any direction.

It is obvious that the disappearance of a favoured direction of vibration should be to make γ increase continuously and not attain a maximum, since the direction at which γ is maximum is the favoured direction. This has been arrived at from consideration of values for γ greater than 0·5 per hour. This may not, and as we shall show later does not, hold for values of γ of the order of about 0·04 per hour. If a favoured direction, however feeble, develops, the several criteria discussed above will enable one to fix its nature very easily without numerous trials having to be carried out. A single trial is sufficient. Holding the thread with pintong jaws (see fig. 4) from the top support, fig. 5 shows to what extent the rotation is

independent of directions N.–S., E.–W., and 25 to 30° N.E.–S.W. The value of θ is slightly low, due to certain corrections for the progress of the apse of the ellipse not having been applied (see later on). In the following figure (fig. 6) the value of the major axis is plotted logarithmically against time. The tangent to the curve gives $k/2$, half the value of the damping constant. The values of the minor axis, as well as γ, are shown. The importance of the criterion

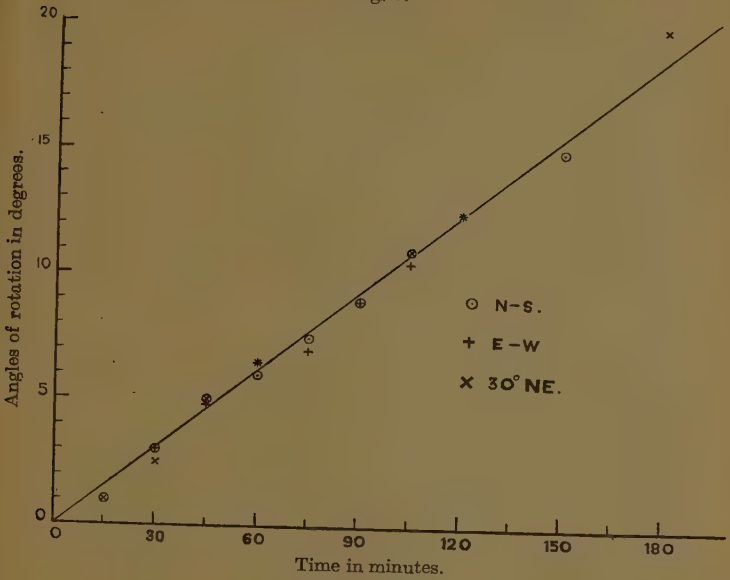
Fig. 4.



Rigid support.

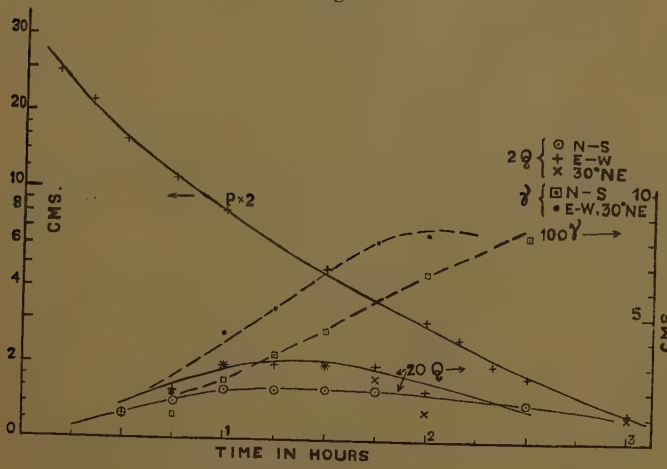
of the favoured direction was not well realized at the time these experiments were carried out, and as also both the major and minor axis become very small towards the end, and no special precautions were taken to strictly exclude any breeze (the effect of breeze becomes appreciable when the amplitudes are very small), the values of γ after 2 hours may be slightly in error. Still they show fairly well that γ has been increasing steadily almost linearly up to 2 hours in all the three cases. Its value is lower in the N.–S. direction. The values in the N.–S. and 30°

Fig. 5.



Rigid support.

Fig. 6.



Amplitudes P, and Q and $\gamma = Q/P$ (rigid support).

N.E. are asymptotically the same in 3 hours, but in the E.W. direction observations were not recorded beyond 2 hours.

The object of the following experiments is not so much to get the most accurate value for the Foucault angle, but more to explore the behaviour and the difficulties and irregularities, and to see what smallest lengths could be used for the pendulum.

In experiments 1–6 (see Table III.) the values for the Foucault angle θ is 6·22, and with apsidal correction it will be 6·6. The correction for the progress of the apse is $\frac{3\pi}{4}\frac{PQ}{l^2}$ radians per period*. In 7–13 the value without apsidal correction is very nearly the correct value. Taking the correction into consideration we get 7·03, a value rather too high. Similarly in 14–18 we get 6·66, but only 6·34 in 19–21 ; again, 6·89 in 22–25 and 6·37 in 26–28 for the smallest length of 96 cm. The average in all the 28 experiments is 6·38+·34=6·72 per hour. How this rather too high a value is obtained is not clear. Probably the nature of the string and the rather rough degree scale used for measuring deflexions have something to do with this.

In the table results of experiments are given for various lengths of the pendulum, using small as well as big bobs. The apsidal correction which is to be added in nearly all cases is given next the measured values for the Foucault angle. Better values are obtained on an average if no apsidal correction is applied. The minor axis gets a maximum value in about 15 to 20 minutes time, using the smaller bob B, but γ is not found to attain a maximum value within the period of observation, namely one hour. At the end of one hour with the small bob B, which has been mostly used, the amplitudes starting with 11 cm. diminishes to about half of 0·3 cm. But the readings of θ could be made accurately even with this small length, because the shadow of the pin-point remains sharp and the position of the major axis easily determined. The angle is highly susceptible to slight faults in fixing, and it requires great practice to fix accurately such that with fixing undisturbed the values of θ are the same in

* Sir G. B. Airy, Phil. Mag. ii. pp. 147–149 (1851). A. Galbraith & S. Haughten, Phil. Mag. ii. pp. 134–139 (1851). A. Thacker, Phil. Mag. ii. p. 159 (1851) ; ii. pp. 412–413 (1851).

orthogonal directions. As mentioned already, with correct fixing the vibration in the ellipse is anti-clockwise, but the Foucault angle of rotation of the major axis is clockwise. These conditions have been fulfilled in all the experiments reported in Table III., except in 2 and 21, where the minor axis has sign negative. But the error is not appreciable, since it has not influenced the Foucault angle to any appreciable extent. The negative sign of the minor axis in 2 and 21 is certainly to be attributed to some slight vibration of the pendulum bob at the moment of release.

Even with the longest pendulum used, namely (53 ft.) 1615 cm., Bunt * was unable to get his pendulum orbit to go anticlockwise in two directions normal to each other. He could get the Foucault angle fairly correctly. With short pendulums of his the deviations are greater The rigidity of the steel wire used must therefore be assumed to have given a most favoured direction. Experiments of Phillips †, too, gave results similar to that of Bunt. A. C. Longdon ‡ has not reported sufficient data to judge the accuracy of his results. It was stated by Bunt that it is impossible to get reproducible results with the use of iron bobs ; he attributed it to the effect of the magnetic field of the earth. We may state that the bobs used by us were of iron. The distributing effects are mostly even, entirely due to faults in fixing the string in the chuck.

We have experimented with brass bobs, and have found that they behave in no way different from iron bobs.

Mathematical Theory.

Let

$$n=\omega \sin \lambda \quad \text{and} \quad p^2=g/l.$$

g is the acceleration due to gravity; then the equations of motion for the bob of the pendulum are

$$\ddot{x}-2n\dot{y}+p^2x=0,$$
$$\ddot{y}+2n\dot{x}+p^2y=0.$$

Introducing a damping term to account for air resistance to the motion, we get

$$\ddot{x}-2n\dot{y}+k\dot{x}+p^2x=0, \quad . \quad . \quad . \quad . \quad (1)$$
$$\ddot{y}+2n\dot{x}+k\dot{y}+p^2y=0. \quad . \quad . \quad . \quad . \quad . \quad (2)$$

* T. G. Bunt, Phil. Mag. ii. pp. 158–159 (1851); ii. pp. 424–427 (1851); iv. pp. 272–275 (1852).

† J. Phillips, Phil. Mag. ii. pp. 150–152 (1851).

‡ A. C. Longdon, Phys. Rev. xiii. pp. 241–258 (1919).

TABLE III.

Length of pendulum.	Bob.	Date 1935.	Serial no.	2 P.	Maximum $20 \times Q$ cm.	t min. for $Q =$ max.	t min. for $k/2 = 1$.	t min. for $k/2 = 2$.	100 γ per hour.	Direction.	θ.	$d\theta$.	Remarks.
cm.													
245	A	11/4	1	44	1	80	28	65	2·4	*m*	6·15°	·26°	
		,,	2	,,	1	85	,,	,,	3·8	*l*	6·2	—·26	
		12/4	3	,,	2	,,	,,	,,	2·4	*m*	6·15	·52	Fixing unaltered.
		,,	4	,,	3	,,	,,	,,	3·8	*l*	6·3	·75	
		14/4	5	,,	3	80	,,	,,	3·8	25 N.E.	6·2	·75	
		15/4	6	,,	3	,,	,,	,,	3·8	,,	6·3	·75	
									3·3		6·22	·38	
231	B	13/9	7	22	2	15	9·5	20	13	*m*	6·4	·27	
	A	14/9	8	,,	1·8	30	25	63	4·5	*m*	6·5	·5	As in 7, but changed bob.
	A	,,	9	,,	3	20	29	66	4·5	*l*	6·5	1·2	Undisturbed.
	A	,,	10	,,	2·2	45	..	..	..	*m*	6·7	·2	Fixing undisturbed.
	B	,,	11	,,	2·5	26	12	25	14	*l*	6·4	·3	Small bob fixed.
	B	17/9	12	,,	2·6	15	11	23	14	*m*	6·6	·36	Refixed bob to thread.
	B	18/9	13	52	2·4	15	9	20	12	*m*	6·6	·6	Untouched.
											6·53	·5	

198	B	26/9	14	22	0·4	15	12	27	1·8	*m*	6·4	·07	Fixing undisturbed.
		,,	15	,,	·5	30	13	28	2·4	*l*	6·5	·07	
		,,	16	,,	1·3	15	13	28	4·0	*m*	6·7	·3	
		,,	17	,,	·5	22	13	26	2·7	*l*	6·6	·09	
		27/9	18	,,	1·0	15	13	26	3·9	*m*	6·4	·2	
											6·52	·14	
167·6	B	2/10	19	,,	2·0	15	11	27	5·0	*m*	6·2	·58	Fixing unaltered.
		,,	20	,,	2·0	,,	,,	,,	5·0	*m*	6·0	·58	
		,,	21	,,	−2·0	,,	,,	,,	8·0	*m*	6·25	−·58	
											6·15	·19	
125·8	B	22/10	22	,,	·2	15	11	23	0·3	*m*	6·25	·06	Fixing unaltered.
		,,	23	,,	1·5	10	10	18	8·0	*l*	6·6	·8	
		,,	24	,,	2·0	15	10	20	5·0	*l*	6·4	1·08	
		25/10	25	,,	0·2	20	11	23	0·5	*m*	6·3	0·06	
											6·39	0·5	
95·3	B	31/10	26	5·0	..	..	..	..	..	*l*	6·4	..	Separate fixings in each case.
		30/11	27	10·4	0·1	..	..	..	..	*m*	6·3	..	
		2/12	28	9·0	..	..	..	..	..	*m*	6·4	..	
											6·37		

Experiments 1–6 were with thin-spun thread of silk without twist; 7–28 with a cotton-string, but with a slightly bigger diameter. Experiments 1–21 were all carried out from a common wooden scaffolding used for fixing the upper support, while 22–28 were supported from the frame of a door between two rooms. In experiments 22 28 the pendulum was very carefully protected from breeze with glass screens. Weight of bob A is 450 gm., steel, and of bob B 150 gm., also of steel. Difficulty of correct fixing increases with diminution in lengths of the pendulum, and heavier bobs allow pendulums to run for greater lengths of time. *m* and *l* refer, respectively, to meridian and latitude.

Let $x+iy=\zeta$, then the equations (1) and (2) reduce to

$$\ddot{\zeta}+(2ni+k)\dot{\zeta}+p^2\zeta=0. \quad . \quad . \quad . \quad . \quad (3)$$

This can be solved by putting

$$\zeta=e^{\mu t}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

and the values of μ are

$$\mu_1=-\frac{k}{2}(1-\alpha)+ip(1-\alpha)=-\frac{k}{2}(1-\alpha)+i\theta_1$$

and

$$\mu_2=-\frac{k}{2}(1+\alpha)-ip(1+\alpha)=-\frac{k}{2}(1+\alpha)-i\theta_2,$$

where $\alpha=n/p$.

Then, substituting in (4),

$$\zeta=A_1e^{-\frac{k}{2}(1-\alpha)t+i\theta_1 t}+A_2e^{-\frac{k}{2}(1+\alpha)t-i\theta_2}$$
$$=\beta_1e^{i\theta_1 t}+\beta_2e^{-i\theta_2 t}. \quad . \quad . \quad (5)$$

Introducing initial conditions, at $t=0$

$\zeta=a$, the amplitude along the x axis, being the initial direction for the swing

$\zeta=0$ along y axis.

We get, on substituting in (6),

$$A_1+A_2=a,$$
$$A_1(p-n)-A_2(p+n)=0.$$

It follows from this that

$$A_1=\frac{a}{2}(1+\alpha)\,;\ A_2=\frac{a}{2}(1-\alpha). \quad . \quad . \quad . \quad (6)$$

The end points of ζ will describe an ellipse with continuously diminishing major axis, *i. e.*, a rosette-ellipse, in which the orbit never closes into a previous orbit. The ratio γ of the minor to the major axis will be given by

$$\gamma=\frac{\beta_1-\beta_2}{\beta_1+\beta_2}=\alpha\left(1+\frac{kt}{2}\right)=\frac{n}{p}\left(1+\frac{kt}{2}\right). \quad . \quad . \quad (7)$$

This agrees with the value calculated by T. J. I'a. Bromwich*

* T. J. I'a. Bromwich, Proc. Lond. Math. Soc. (2) xiii. p. 222 (1914). The review of this paper by Auerbach and Hort in *Handb. d. Physik u. Tech. Mechanik.* ii. p. 665 (1930) is erroneous : α should be $\sqrt{g/l}$, and 2·6 mm. is the distance moved by the apse in the outer circle.

if we neglect from (7) effect due to air damping and put $k=0$.

The values of $\frac{kt}{2}$ are shown in Table III., and the other values are $\omega=7{\cdot}3.10^{-5}$; $\sin\lambda=0{\cdot}427$; $p=\sqrt{\frac{g}{l}}=2{\cdot}0$; $n=3{\cdot}12.10^{-5}$ for pendulum experiments 1–6 in Table III.

Substituting in (7), we get the absurdly low value for γ, namely $4{\cdot}68.10^{-5}$ per hour, whereas the value actually observed is about $3{\cdot}3.10^{-2}$ per hour. This experimental value, 705 times greater than the theoretical one, cannot be due to experimental error, as all previous investigators thought it was, arising somehow from rigidity of the wires they used, faulty fixings, etc., Supposing it were, how is it that the Foucault angle of rotation is not affected to any appreciable extent ?

We shall now investigate the implications of (7):

$$\gamma=\alpha\left(1+\frac{kt}{2}\right)=\frac{Q}{P}=\frac{Q}{P_0e^{-\frac{kt}{2}}}\simeq\frac{Q}{P_0}\left(1+\frac{kt}{2}\right).$$

Therefore

$$\alpha=\frac{Q}{P_0}\quad\text{or}\quad Q=\alpha P_0=1{\cdot}56.10^{-5}.22=3{\cdot}44.10^{-4}\text{ cm.}\qquad\ldots\quad(8)$$

From (8) it follows, as both α and P_0 are constants, that Q will be constant in spite of the fact that the major axis will be continuously and rapidly diminishing, due to air damping. Its magnitude comes out to be such that it would require a powerful microscope to detect it. In contrast to this, the value for Q experimentally observed in almost every case gives that, starting from zero, it rises to its maximum value of about 0·15 cm. in about 82 minutes in experiments 1–6, and then gradually diminishes. With smaller pendulum lengths the maximum is reached much earlier. How are these to be explained ?

As Q is not constant, let us assume that dQ/dt at any time is proportional to the value of P at that time. Taking into account the damping of the minor axis as well as that of the major axis we have

$$dQ/dt=A\alpha P-\tfrac{1}{2}k_2Q$$

and

$$dP/dt=-\tfrac{1}{2}k_1P.$$

These equations have the solution

$$Q=\frac{2A\alpha P_0}{k_2-k_1}(e^{-k_1T/2}-e^{-k_2T/2}), \quad . \quad . \quad . \quad (10)$$

satisfying the initial condition that Q=0 at T=0. Q rises rapidly from its zero value at time zero, attains a maximum, and then falls off gradually, in agreement with experimental values, curves for Q in fig. 6.

The equation (10) is analogous to that which holds for the growth of an intermediate product in a radioactive change. While Q grows with time its decay is also proceeding at a definite rate (see equation (5), p. 11, and fig. 2, curve B, in 'Radiations from Radioactive Substances,' by Lord Rutherford, Chadwick, and Ellis, or chapter on Radioactivity in 'Electricity and Magnetism,' by S. G. Starling).

Maximum of Q.

Differentiating Q in (10), we get the condition for $Q_{max.}$, namely,

$$-\frac{k_1}{2}e^{-\frac{k_1T}{2}}+\frac{k_2}{2}e^{-\frac{k_2T}{2}}=0,$$

or

$$\frac{k_1}{k_2}=e^{(k_1-k_2)T/2}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

By plotting the logarithm of the major axis against time (see fig. 6), and then plotting from this the tangents of these curves against the amplitude, the values of $\frac{k}{2}$ can be determined as a function of time. For experiments 1–6 the following table holds good :—

TABLE IV.

Time (min.).	Amplitude 2 P (in cm.).	tan θ.	$\frac{k}{2}\times 60$.
22	20	1·63/117	·0321
50	10	1·49/155	·0221
85	5·3	1·37/173	·0182
127	2·5	1·23/190	·0149
187	1·02	1·0/204	·0113

The values of both $\frac{k_1}{2}$ and $\frac{k_2}{2}$ have to be extrapolated for *average* amplitude. Let us take $\frac{k_1}{2}$ as ·025 and $\frac{k_2}{2}$ as ·005. Here, in choosing k_1 and k_2 we may be guided by the fact that T_m is somewhere near about 80 as experimentally known. We get T_m for Q max. as 80 minutes, in close agreement with experimental value of 82 minutes. Taking $\frac{k_1}{2}$ as 0·02, $\frac{k_2}{2}$ as ·005, we get T_m as 95, also not far from the observed value.

Values of γ.

From (10) it follows at once that

$$\gamma = \frac{2A\alpha}{k_1 - k_2}(e^{(k_1 - k_2)T/2} - 1). \quad . \quad . \quad . \quad . \quad (12)$$

For small values of T (12) becomes

$$\gamma = A\alpha T \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (12\,a)$$

$$= \bar{\gamma} T. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (13)$$

It follows from this that initially γ must increase directly with time. This is found to be so in the actual experiments; see the lines representing γ in fig. 6. The lines are straight for the major portion of the time during which the observations could be made.

Value and Nature of A.

Using the equations (13) and (12 *a*), we can calculate the values of A, knowing the experimentally determined values of γ, as the values of α in terms of n and p are also known; n is a constant depending only on the geographical latitude of the place of observation, and $p = \sqrt{g/l}$ is determined from the value of the gravitational acceleration and the lengths of the pendulum in the individual cases.

The following table gives the values of A :—

TABLE V.

The values of A, of equation (12 *a*).

Experiments.	Value of A (average).	Remarks.
1–6	0·59	Silk-string, bob A, $l = 245$ cm.
8–9	0·82	Cotton-string, bob A, $l = 231$ cm.
7, 11–13	2·4	Cotton-string, bob B, $l = 231$ cm.
14–18	0·59	Cotton-string, bob B, $l = 198$ cm.
19–20	1·1	Cotton-string, bob B, $l = 167{\cdot}6$ cm.
22–25	0·87	Cotton-string, bob B, $l = 125{\cdot}8$ cm.
	1·01	weighted average value of A.

We reach the important conclusion that the value of A is *unity*, in spite of the fact that it varies within fairly wide limits. This means that the variation of γ with time, or $\dot{\gamma}$, is equal to α itself; therefore it is a measure of α. *Hence the rotation of the earth can be calculated from* $\dot{\gamma}$, *and so we have in this a new method for its determination.* As $\dot{\gamma}$ has dimension of -1 in T, while α has zero dimension, it is proposed to retain A in equation 12 (*a*). More accurate determination of the value of A is necessary, and we hope to do it shortly.

Will γ *reach a Maximum Value* ?

From (12), it is obvious that γ will rise exponentially; but this need not be strictly true, as both k_1 and k_2 vary slightly with the magnitudes of the major and minor axes. Since the minor axis is small we shall assume k_2 as constant, but instead of $k_1/2$ we shall substitute the value

$$\tfrac{1}{2}k_1(1-a\mathrm{T})$$

to express the diminution in damping of the major axis as its length becomes smaller and smaller with time.

Substituting the new damping term for $k_1/2$ in (12), we get

$$\gamma \simeq e^{(k_1-k_2)T/2-k_1\alpha T^2/2}.$$

From this we get the condition for γ maximum, namely

$$(k_1-k_2)/k_1=2\alpha T_{max.} \simeq 0{\cdot}8$$

(on substituting the actual values of k_1 and k_2). The value of a can be calculated from Table IV. as 0·1 in 30 minutes, the value for $T_{max.}$ comes out as 20 hours, in experiments 1–6. The experimental value would denote a maximum, but would be difficult to determine, since the accuracy in the measurement of the minor axis beyond $2\frac{1}{2}$ hours is very doubtful and has not been observed for a period longer than 3 hours.

We could write down the final expressions for Q, γ, and $\bar{\gamma}$ as follows :—

$$Q=\frac{2\alpha P_0 A}{k_2-k_1}(e^{-k_1 t/2}-e^{-k_2 t/2}). \quad . \quad . \quad . \quad (10')$$

$$\gamma=\frac{2\alpha A}{k_1-k_2}e^{(k_1-k_2)t/2}-1). \quad . \quad . \quad . \quad . \quad (12')$$

$$\bar{\gamma}=\alpha A. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (14')$$

Here $\alpha=n/p=\omega \sin \lambda \sqrt{l/g}$; A is a constant, nearly equal to one, and has dimension of -1 in T.

Summary.

The Foucault angles or rotation of the plane of oscillation of pendulums are studied for lengths of the same between 245 cm. and 95 cm. Threads of cotton as well as of silk were used for suspensions. The weights of the bobs were 450 and 150 gm. and they were made of steel. Control experiments were also done with bobs of brass of the same size as that of the steel bob of 450 gm. The suspension of thread is shown to overcome many difficulties—difficulties which increase as the length of the pendulum is decreased. The errors arising from the supports of the suspension threads have been studied by a special type of knot suspension. This enables one to interpret many of the older results in the literature. By a careful elimination of the errors the second order effect in the Foucault's pendulum with rigid supports, namely, the ellipticity of the orbits of the bob, is shown to be

capable of investigation. The theory of the Foucault's pendulum has been extended beyond that given by T. J. I'a. Bromwich by taking into consideration the effect of damping due to air. Further modification is found to be necessary, and has been made and is given in the expressions (10′), (12′), and (14′). These bring into expression the fact that the ellipticity, defined as the ratio of minor to major axis observed, is very much greater than can be accounted for by damping due to resistance of air. It is believed that a new method of measuring the rotation of the earth is found in expression (14′), with, however, certain difficulties, discussed on p. 86, to be overcome.

VII. *The Potential Difference at an Air-Water Interface.* *By* J. A. Chalmers, *M.A.*, *Ph.D.* (*Cantab.*), *Lecturer in Physics, Durham University* (*Durham Division*), *and* F. Pasquill, *B.Sc.*, *Pemberton Research Fellow, University College, Durham* *.

1. *Introduction.*

THERE are a number of experiments which have demonstrated the existence of a potential difference at an air-water interface, but so far it has not been found possible to obtain an accurate estimate of this potential difference.

The experiments concerning the potential difference may be divided into two main categories: in the first place there are experiments on the cataphoresis of air-bubbles in water (Alty [1], McTaggart [2]); by the application of the theoretical laws of cataphoresis (Helmholtz [3]), McTaggart calculated the potential difference to be ·055 volts, but this assumed a constant velocity of cataphoresis for all sizes of bubble, and Alty showed that this is true only over a very limited range of bubble sizes. It does not seem possible to use Alty's results to give a value of the potential difference at the air-water interface. These experiments concern only what is termed by Freundlich [4] the "electrokinetic potential difference," *i. e.*, the potential difference across the boundary at which the motion occurs, and not the total potential

* Communicated by the Authors.

difference from the interior of the liquid to the air. Experiments, such as those of Lenard [5], on the splashing of drops indicate that a negative charge may be more easily removed from a water surface than a positive charge, and hence we can infer the existence of a double layer of charge, but no indication is given of the corresponding potential difference.

On the other hand, there are several experimental methods which concern the air-water potential difference and the influence of dissolved substances in the water, but, at the same time, at least two other potential differences are involved. The experiments of Ayrton and Perry [6] on the Volta potential difference between a metal and water give the sum of the potential differences at the air-water, water-metal, and metal-air boundaries. The method of Kenrick [7], Frumkin [8], and Kamienski and Goslawski [9] consists in measuring the potential difference produced when a liquid jet breaks into drops inside a hollow tube down the inner surface of which is flowing another liquid ; here we have two liquid-air interfaces and two metal-liquid interfaces, assuming the same metal is used for all connexions. In the method of Guyot [10], Frumkin [11], and Schulman and Rideal [12], ionization of the air allows that a platinum point shall acquire the potential of the air just outside a liquid, and this is compared with that in the liquid ; here we have the liquid-air interface, the air-platinum interface, and a platinum-liquid interface, provided that platinum is used for all connexions, otherwise we have also a metal-metal interface.

In obtaining single potential differences at a metal-liquid boundary, various methods can be used to eliminate all but one of the boundary potential differences, and so to allow this one to be measured, for example, by the capillary electrometer [13], the dropping electrode [11], or the method of Garrison [15]. But none of these methods seems to be available to eliminate potential differences in the case of air-liquid boundaries.

However, it is clear that the mechanism of the production of a potential difference at an air-liquid boundary must lie in the orientation of the liquid dipole molecules at the surface and in the selective adsorption of ions of one sign on to the air surface. Therefore if we can introduce some other surfaces which will, by their nature,

orientate the molecules and adsorb the ions elsewhere than at the air surface, we may prevent the production of the potential difference at the air surface; we can, at the same time, avoid providing any fresh potential difference if the surfaces used are such as to give random distribution of the orientations of the molecules and the positions of the adsorbed ions. It is suggested that such conditions are fulfilled by using some absorbent material such as filter paper soaked in water, for there will be many surfaces in random directions in the filter paper. By using a filter paper-air boundary, at which, according to the above argument, there is no potential difference, it is suggested that the potential difference at an air-water interface can be obtained.

2. *Experimental Method.*

The method used for measurement is to form water drops inside a hollow tube which is lined with wet filter paper (fig. 1); the drops are then collected and the charge on each is measured by means of a two stage valve bridge amplifier. Different potentials are applied between the dropper and the tube, and the variations of drop charge with the applied potential is plotted, giving as intercept the potential that must be applied so that the drop carries no charge (fig. 2).

On the analogy of Kelvin's water-dropper [16], we see that the charge carried down depends on the potential difference across the air space between the drop and the tube, for this is associated with lines of force which end on the drop. Therefore the applied potential for which the drop carries no charge is just opposite to the total natural potential difference of the system.

In most of the experiments performed both dropper and tube consisted of copper, but in some cases, as described later, the dropper is of glass with a zinc rod dipping into it, while the hollow tube is also of zinc.

3. *Experimental Results and their Interpretation.*

In the following we shall use abbreviations to denote the potential differences at the interfaces, such as V_{ca} for the potential difference at the copper-air interface, in each case taking the first named as positive; z, w, and f refer to zinc, water, and filter paper respectively.

Fig. 1.

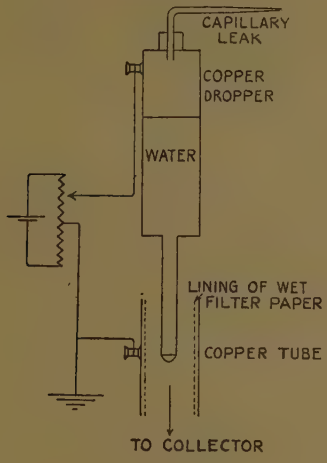
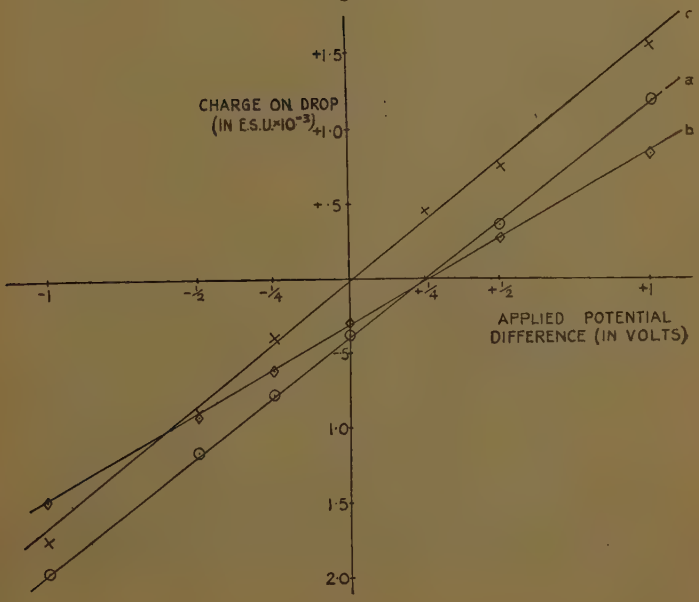

CAPILLARY LEAK
COPPER DROPPER
WATER
LINING OF WET FILTER PAPER
COPPER TUBE
TO COLLECTOR

Fig. 2.


CHARGE ON DROP (IN E.S.U.×10⁻³)
+1·5
+1·0
+·5
-·5
-1·0
-1·5
-2·0
-1
-½
-¼
+¼
+½
+1
APPLIED POTENTIAL DIFFERENCE (IN VOLTS)
c
a
b

Using the experimental method described above, it is found that the drops carry zero charge if the dropper is at a potential of ·26 volts above that of the tube. This means that the total natural potential difference is −·26 volts, *i. e.*, that

$$V_{wc}+V_{cf}+V_{fa}+V_{aw}=-\cdot 26 \text{ volts.} \quad . \quad . \quad (1)$$

If, in the external connexion between the dropper and the tube we insert wet filter paper and water, it is found that there is no change in the natural potential difference. Therefore,

$$V_{wc}+V_{cw}+V_{wf}+V_{fc}+V_{cf}+V_{fa}+V_{aw}=-\cdot 26 \text{ volts,} \quad (2)$$

or

$$V_{wf}+V_{fa}+V_{aw}=-\cdot 26 \text{ volts.} \quad . \quad . \quad . \quad . \quad (3)$$

From (1) and (2) we deduce

$$V_{wc}+V_{cf}+V_{fw}=0. \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

Now if we use a dropper with zinc and a tube of zinc, we get the same intercept, so that

$$V_{wz}+V_{zf}+V_{fa}+V_{aw}=-\cdot 26 \text{ volts.} \quad . \quad . \quad . \quad (5)$$

From (5) and (1) we get

$$V_{wc}+V_{cf}=V_{wz}+V_{zf}. \quad . \quad . \quad . \quad . \quad . \quad (6)$$

We know that V_{wc} and V_{wz} are not the same, and hence the most reasonable deduction from (6) is that

$$V_{wc}=-V_{cf} \quad \text{and} \quad V_{wz}=-V_{zf}. \quad . \quad . \quad . \quad (7)$$

Applying this to equation (4) we see that

$$V_{fw}=0. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

and now in (3) we have $V_{fa}=0$ from the earlier assumption, so that we deduce the final result that

$$V_{aw}=-\cdot 26 \text{ volts,} \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

i. e., there is a potential difference of ·26 volts at an air-water interface, the water being positive.

It was also found that the same result occurs whether ordinary tap-water or distilled water was used. When hot water was put into the dropper the same result occurred.

4. *Investigation of Other Possible Effects.*

The fact that the drop acquires a charge when there is no applied potential difference might be ascribed to many other causes than that considered ; an exhaustive study was therefore made of the conditions, the following points being investigated.

(a) *The Medium holding the Water inside the Tube.*

The potential difference observed may be due to the filter paper itself, so other materials were used in its place. The following materials gave results identical with that for filter paper : linen bandage, cotton wool, silk, velvet, glass wool, asbestos paper, asbestos wool, and barium sulphate precipitate. The only materials not giving the same effect were chamois leather and red felt, and the latter agreed with filter paper after boiling to remove dye. Since the only property common to all the materials used is that of holding water, it is clear that the effect is due to the water and not to the absorbent material.

(b) *Size of Drop.*

If the effect found is really due to a potential difference, the alteration in size of the drop should not alter the potential difference found as intercept. This is found to be so, as indicated in fig. 2 (*b*), which shows results for a smaller drop.

(c) *Type of Nozzle.*

If the charge on the drop were due to friction at the nozzle, the charge would depend on the type of nozzle used. Experiments were carried out in which the water drops from a glass tube, the metal in the circuit being zinc, not copper, and the results obtained agreed with those for the copper dropper. No difference was found when the drops fell from a strip of zinc. Further, alterations in the shape of the copper nozzle do not produce any effect on the results.

(d) *Height of Fall of Drop.*

Variation of the height through which the drop falls does not alter the results, showing that the drop does not acquire an appreciable charge while falling.

(e) *Time of Formation of Drop.*

The time taken for successive drops to form was varied by altering the capillary leak (see fig. 1), and it was found that for all times between three seconds and three minutes the results obtained were identical.

It might be thought that the charge on the drop was produced in a way analogous to that occurring in the dropping electrode [14], where the drops have not time to acquire the surface potential difference. But it is clearly more reasonable to suppose that the whole of the potential difference is set up in less than three seconds rather than that it has not started to appear in three minutes.

(f) *Charge passing to the Dropper.*

If the dropper and shield are connected with the collector it is found that no resultant charge passes as the drop falls; this verifies that no charge is picked up from the air by the falling drop.

If the dropper alone is connected to the measuring system it is found that a charge flows to the dropper as the drop falls, and this is about two-thirds of the charge on the drop and of opposite sign. The charge passing to the dropper becomes zero for an applied potential difference of ·26 volts, so that this must be the potential required to give the charge to the newly formed drop. This verifies that no appreciable time is required for the setting-up of the potential difference.

The recurrence of the ·26 volts proves beyond doubt that this represents a real potential difference.

(g) *Running Water in Tube.*

If the filter paper inside the tube is replaced by running water (as in the method of Kenrick [7] etc.) it is found that zero charge is obtained for zero applied potential difference (fig. 2 (*c*)). This verifies that the value of ·26 volts is due to the filter paper, and not to any effects connected with the drops.

(h) " *Water-Break* " *Surface in Tube.*

Finally the inner surface of the tube was " degreased " by trichlorethylene vapour and sodium metasilicate solution, so that it would carry the " water-break "

surface, and this stationary water-surface gave the same result as the running-water surface ; this verifies that the effect is due to the filter paper, not to the drops ; also, since the " water-break " surface persists for some time and gives the same results, it proves the absence of any effect due to the time taken to build up the potential difference.

5. *Conclusion.*

Without any assumptions at all, the experiments performed have proved that there is a value of $-\cdot 26$ volts as the sum of V_{wf}, V_{fa}, and V_{aw}.

This has been interpreted as meaning that V_{wa} is $\cdot 26$ volts, and it is difficult to conceive any other interpretation that will account for the results.

This represents the total potential difference across the interface, or at any rate that part of it which is eliminated by means of the filter paper. There is no doubt that this is greater than the electrokinetic potential difference that is of importance in the cataphoresis experiments. Alty [1] has found that the velocity of cataphoresis depends on the purity of the water and also that the charge on the air bubbles takes an appreciable time to attain its final value ; it is, however, not surprising that there is no appearance of these phenomena in the present experiments, for the magnitude of the electrokinetic potential difference may well be much smaller than the potential differences appreciable in the present experiments.

It may be suggested that the $\cdot 26$ volts found is due to the orientation of the water molecules in the surface, and that the differences found by Frumkin [8], [11] etc. may be due to adsorbed ions. Electrokinetic phenomena must be due entirely to the adsorbed ions, for a potential difference due to a non-separable double layer (such as orientated molecules) cannot give rise to electrokinetic phenomena, which depend on the relative motion of the two portions of the double layer. A consideration of the probable magnitudes of the effects due to orientated molecules and to adsorbed ions suggests that the former may well be much greater.

If we consider the water molecule to have a dipole moment of $1\cdot 9 \times 10^{-18}$ e.s.u. [17], then we have $\frac{4 \times 1\cdot 9 \times 10^{-18}}{A}$

$= \cdot 26$ volts $= \frac{\cdot 26}{300}$ where A is the average area occupied by an orientated molecule in the surface, giving $A = 2 \cdot 8 \times 10^{-14}$ sq. cm., corresponding to about 1 water molecule in 25 on the surface being completely orientated, or a larger number partially orientated, a reasonable result.

6. *Summary.*

Experiments described show a value of ·26 volts for the sum of the potential differences water-air, air-wet filter paper, and filter paper-water. Subsidiary experiments are described, and it is suggested that this is a manifestation of the potential difference at the water-air boundary, probably due to a double layer consisting of orientated water molecules at the surface.

References.

(1) Alty, Proc. Roy. Soc. cvi. p. 316 (1924); cxxii. p. 633 (1929).
(2) McTaggart, Phil. Mag. xxvii. p. 297, xxviii. p. 367 (1914); xliv. p. 386 (1922).
(3) Helmholtz, *Ann. der Phys.* vii. p. 337 (1879).
(4) Freundlich, 'New Conceptions in Colloid Chemistry,' p. 18 *et seq.*
(5) Lenard, *Wied. Ann.* xlvi. p. 584.
(6) Ayrton & Perry, Phil. Trans. Roy. Soc. clxxi. p. 15 (1880).
(7) Kenrick, *Zeit. Phys. Chem.* xix. p. 625 (1896).
(8) Frumkin, *Zeit. Phys. Chem.* cix. p. 34; cxi. p. 190 (1924).
(9) Kamienski & Goslawski, *Bull. Int. de l'Acad. Polon. des Sci. et Lett.* p. 309 (1933).
(10) Guyot, *Ann. Phys.* ii. p. 516 (1924).
(11) Frumkin, *Zeit. Phys. Chem.* cxv. p. 485 (1925).
(12) Schulman and Rideal, Proc. Roy. Soc. cxxx. p. 259 (1931).
(13) Lippmann, *Ann. der Phys.* v. p. 494 (1875).
(14) Palmaer, *Zeit. Phys. Chem.* lix. p. 129 (1907).
(15) Garrison, J. Amer. Chem. Soc. xlv. p. 37 (1923).
(16) Kelvin, 'Nature,' li. p. 495 (1895).
(17) Debye, 'Polar Molecules,' p. 40.

VIII. *The Determination of Young's Modulus by Flexural Vibration.* *By* G. Grime, *M.Sc., and* J. E. Eaton *.

[Plate I.]

Introduction.

IN a previous paper [1], an account was given of a method of determining Young's Modulus by observation of the natural frequency of longitudinal vibration of a prism. The method was developed for the examination of building

* Communicated by A. F. Dufton, M.A., D.I.C.

materials such as tiles, slates, and bricks, which are troublesome to deal with by methods involving measurement of a change of length. Owing to the difficulty of exciting and recording high frequencies, it is not convenient to record the longitudinal vibrations of specimens less than about six inches in length, and an alternative method, in which Young's Modulus is determined from the frequency of flexural vibration, has been developed. It is found that the method is not only applicable to specimens of brick or tile, but is very convenient also for materials of high internal friction such as pitch and asphalt.

Theoretical.

The frequency of free flexural vibration of a prism is [2]

$$n=\left(\frac{\mathrm{E}}{\rho}\right)^{\frac{1}{2}}\left\{\frac{m^2\mathrm{K}}{2\pi l^2}\right\},$$

where E is the Young's Modulus and ρ the density of the material; l the length of the prism and k the radius of gyration of the cross-section about its intersection with the neutral surface; and where m is a number to which the appropriate value has to be given.

When the bar is free at both ends and vibrates at its fundamental frequency m has the value 4·73 and

$$\mathrm{E}=\frac{4\pi^2l^4n^2\rho}{m^4\mathrm{K}^2}.$$

The frequency n of the above expression is that of undamped flexural vibration. For substances such as asphalt, in which vibration is highly damped, it becomes necessary to determine whether the actual frequency of oscillation differs from the theoretical value for undamped vibration by an amount sufficient to give rise to serious error. The amplitude of a damped simple harmonic vibration is

$$a=a_0e^{-kt}\sin(4\pi^2n^2-k^2)^{\frac{1}{2}}t,$$

where a_0 is the amplitude at the time $t=0$ and k is a constant depending on the internal friction. The actual frequency, therefore, is $\{n^2-(k/2\pi)^2\}^{\frac{1}{2}}$ and is not significantly different from n unless k is nearly as great as n. k is readily found by measuring the decrement, and it will be shown later that, with the accuracy attainable, the difference between the two frequencies is of no significance.

Experimental.

With the exception of the piezo-electric detector the recording apparatus has been fully described in the previous paper.

The specimen, in the form of a bar of uniform cross-section, is supported on two pads of sponge rubber placed at the two nodes of the fundamental vibration for a bar free at both ends, that is, at a distance of 0·224 l from either end. A small piezo-electric quartz crystal cemented to the specimen half-way along its length forms the detecting device. The vibration is excited by a blow from a light hammer, and the alternating voltage set up is amplified and recorded with a cathode-ray oscillograph. A single-traverse time-base, triggered by the stroke of the hammer, spreads out the vibration pattern, which is photographed with a stationary plate camera. The frequency is determined by comparison with a scale of milliseconds recorded on the same plate.

It is important to ensure that the mode of vibration is correctly identified. A very simple test with a specimen mounted as described shows that it vibrates in the fundamental mode for a bar free at both ends; the fluorescent spot is observed visually, the time-base being disconnected, and the position at which the blow is struck is gradually shifted along the bar; it is found that the amplitude of the vibration falls to zero as each of the two nodal points is approached, proving that the bar is vibrating in its fundamental node. The nodal points, indeed, can be located with considerable accuracy by this method.

The accuracy of the absolute determination of Young's Modulus by the method of flexural vibration depends chiefly on the accuracy with which the section of the prism can be measured. It is estimated that with a rectangular prism measuring 4 in. × 1 in. × 0·5 in. the probable error is approximately 2 per cent.

Pl. I. shows examples of the records of flexural vibration. Fig. (*a*) is a record from a specimen of wallboard measuring 9·80 in. × 1·77 in. × 0·20 in. The frequency is 176 per second, the density is 52·5 lb. per cub. ft., and the calculated value of Young's Modulus is $0{\cdot}54.10^6$ lb. per sq. in. Figs. (*b*) and (*c*) (Pl. I.) are records of the vibration of specimens of rock asphalt and of an asphalt mastic, a road carpeting material. It will be noted that the vibrations of the two last-named materials are very highly

damped. This is characteristic of substances such as pitch, asphalt, and bituminous materials, many of which can, nevertheless, be induced to yield flexural vibration records. It has been shown above that the period of a highly damped system is slightly greater than that of the same system undamped. In practice, however, the difference is negligible ; in the extreme case of fig. 1 (*c*) the damping only reduces the frequency from 1239 to 1230 per second.

A Table is given below of values of Young's Modulus determined by flexural vibration. The materials were chosen to illustrate the scope of the method.

Material.	Young's Modulus (lb. per sq. in.) at 18° C.
Lead (commercial)	$2{\cdot}63.10^6$
Ebonite	$0{\cdot}71.10^6$
Wall-board	$0{\cdot}54.10^6$
Paraffin wax	$0{\cdot}27.10^6$
Rock asphalt	$0{\cdot}09.10^6$
Asphalt mastic	$1{\cdot}7\ .10^6$

References.

(1) Grime, G., Phil. Mag. (7) xx. pp. 304–310 (1935).
(2) Barton, E. H., 'A Text book on Sound,' p. 281. Macmillan & Co., Ltd. (1926).

The Building Research Station,
Garston, May 1936.

IX. *The Inductance of Iron-cored Coils carrying Direct Current.* *By* G. F. Partridge, *B.Sc., F.Inst.P.**

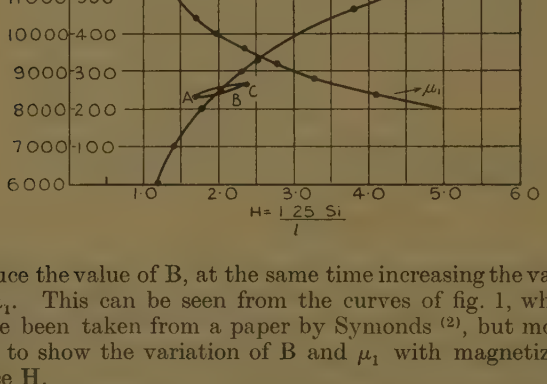
THE inductance of an iron-cored coil to alternating current, when the alternating current is superimposed upon a direct current is, in general, very much smaller than in the absence of direct current. This has been shown to be due to the fact that the permeability of the iron to alternating current is considerably reduced by the presence of a direct current. When alternating current only is passing through the coil the permeability μ

* Communicated by the Author.

is equal to the ratio B/H where B is the flux density in the iron (see fig. 1). When direct current is also passing there is a mean flux density B produced by this current, but the variation of flux density about this mean value due to the alternating current follows a small loop, such as ABC in fig. 1, and the permeability in this case is determined by the slope of this loop. The slope of the loop $\Delta B/\Delta H$ is called the incremental permeability μ_1, and as may be seen from the figure is much less than B/H at the mean value B.

It has been pointed out by Hanna [1] that the inductance under these conditions can be raised by introducing an air-gap into the iron circuit. The effect of the gap is to

Fig. 1.



reduce the value of B, at the same time increasing the value of μ_1. This can be seen from the curves of fig. 1, which have been taken from a paper by Symonds [2], but modified to show the variation of B and μ_1 with magnetizing force H.

The flux density B and the inductance L under these conditions are given by

$$B = \frac{1\cdot25\ N.I}{\frac{l_i}{\mu} + l_a} = \frac{1\cdot25\ N.I}{l_i\left(\frac{1}{\mu} + \alpha\right)} \text{ lines cm.}^2 \quad . \quad . \quad (1)$$

$$L = \frac{1\cdot25\ N^2A \times 10^{-8}}{\frac{l_i}{\mu_1} + l_a} = \frac{1\cdot25\ N^2A \times 10^{-8}}{l_i\left(\frac{1}{\mu_1} + \alpha\right)} \text{ henries} . \quad (2)$$

where N = Number of turns.

I = Direct current in the coil in amperes.

l_i = Length of iron path in cm.

α = Gap ratio $= \dfrac{\text{length of gap}}{\text{length of iron}} = \dfrac{l_a}{l_i}$.

A = Area of cross-section of iron in cm.^{2}.

From these two expressions and the curves showing the variation of B and μ_1 with H, design curves have been worked out [1], [2], from which it appears that there is an optimum gap ratio for any particular winding and shape of iron circuit.

In order to design a coil of this kind accurately it is essential to have the correct length of gap. It has been shown by the author [3] that with gaps of the order required there is always leakage flux which has the effect of reducing the gap length, and this leakage flux depends upon the shape of the iron circuit. It has been proposed by Symonds [2] that tests be made on two samples of the iron with different lengths of gap to determine the relationship between actual length of gap and effective length of gap. It is much more satisfactory, however, to be able to calculate this, and it is proposed to show how this can be done with sufficient accuracy for the purpose.

Variation of Inductance with Length of Gap.

Before proceeding to do this we will see how the inductance varies with length of gap for a particular case. For the choke coil of dimensions shown in the insert of fig. 2, it has been found from Symonds's curves that the optimum gap ratio is about 0·0025, taking a variation of flux ΔB of 89 lines R.M.S., the number of turns being 7000 and the direct current 60 milliamperes.

From the figure A = 6·5 cm.^{2}, l_i = 24 cm., and inserting these values in equations (1) and (2) we get

$$B = 21{\cdot}9/\left(\frac{1}{\mu} + \alpha\right) \text{ lines/cm.}^2$$

$$L = 0{\cdot}165/\left(\frac{1}{\mu_1} + \alpha\right) \text{ henries.}$$

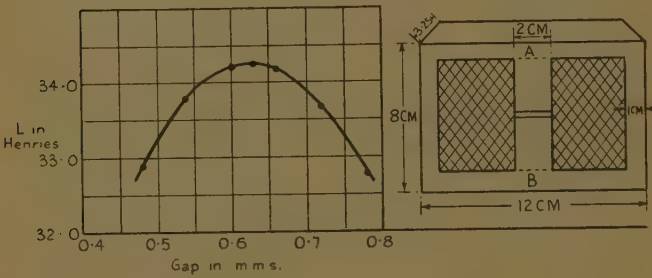
From these expressions the following table of values has been calculated for gap ratios ranging from 0·0020 to 0·00325, and the results are shown graphically in fig. 2.

It will be seen that a variation of ± 10 per cent. in length of gap about the optimum value only produces a 1 per cent. change in inductance, so that if the effective gap can be estimated within this tolerance it will be sufficient for practical purposes.

Calculation of Effective Gap.

As far as leakage is concerned the only part of the magnetic circuit which need be considered is the iron core AB. In actual practice the winding cannot be made to fit

Fig. 2.



Variations of L with Length of Gap.

$\alpha \times 10^4$.	20.	22·5.	25.	26·2	27·5	30.	32·5.
l in mm.	0·48	0·54	0·60	0·63	0·66	0·72	0·78
B	9500	8700	7980	7660	7360	6800	6300
H	2·6	2·12	1·74	1·61	1·50	1·32	1·21
μ_1	330	380	430	455	480	530	560
L Henries	32·9	33·8	34·2	34·25	34·2	33·7	32·8

exactly around the iron core, so that there is always an air space. In addition there is the space occupied by the insulation. This space is sufficient for magnetic leakage to take place across the gap.

Now the magnetomotive force between A and B drives the magnetic flux which passes through the two iron limbs and across the gap. The flux across the gap is therefore proportional to this force divided by the reluctance of the gap, since that of the iron limbs can be ignored, *i. e.*, it is inversely proportional to l_a/A, the latter being a measure of the reluctance of the gap, where l_a is the

length of the gap and A is the area of the gap. We see, therefore, that the flux across the gap is directly proportional to A/l_a.

In addition, the magnetomotive force drives the leakage flux from any particular element of area A_1 on one limb through the space surrounding the core to a similar element of area on the other limb, so that, as before, the leakage flux between these elements is proportional to A_1/z_1 where l_1 is the particular length of the air-path.

The total leakage flux is therefore proportional to the sum of a series of values of A_1/l_1 as we pass along the limbs. The value of the sum of these values of A_1/l_1 can be determined by integration [3]. Thus we find that the leakage flux is proportional to $P/2.\log_e S/l_a$, where S is the length AB (see fig. 2) and P is the perimeter of the iron core.

Again, if we imagine a gap of length l_e and *no leakage* the total flux would be proportional to A/l_e.

If the flux obtained with this imaginary gap of length l_e is the same as the flux found with the actual gap l_a plus the leakage, then l_e is the effective length of gap, so that flux with gap l_e = flux with gap l_a + leakage flux,

or
$$\frac{A}{l_e} = \frac{A}{l_a} + \frac{P}{2} \cdot \log_e \frac{S}{l_a},$$

or
$$\frac{A}{l_a} = \frac{A}{l_e} - \frac{P}{2} \cdot \log_e \frac{S}{l_a},$$

from which we can calculate the actual gap l_a required to be equivalent to an effective gap of length l_e.

It is usually sufficiently accurate if we put l_e for l_a in the last term,
so that
$$\frac{A}{l_a} = \frac{A}{l_e} - \frac{P}{2} \cdot \log_e \frac{S}{l_e}$$

and
$$\frac{1}{l_a} = \frac{1}{l_e} - \frac{P}{2A} \cdot \log_e \frac{S}{l_e}, \quad . \quad . \quad . \quad . \quad . \quad (3)$$

or
$$\frac{1}{l_a} = \frac{1}{l_e} - \frac{1}{l_s}$$

where
$$l_s = \frac{2A}{P} \Big/ \log_e \frac{S}{l_e}.$$

The expression $\frac{1}{l_a} = \frac{1}{l_e} - \frac{1}{l_s}$ will be seen to resemble that for resistances in parallel, so that l_s may be regarded as the length of a shunt gap produced by leakage.

In the present case we require a gap of effective length 0·063 cm. (see table). Also A=6·5 sq. cm., P=10·5 cm., and S=6 cm. Putting these values in equation (3) above we find that l_a=0·082 cm., so that we see the actual gap must be about 30 per cent. longer in order to produce the required effective gap.

Measurements have been made with an experimental coil, and it has been found that the above expression for l_a is sufficiently accurate to give the inductance within a few per cent.

Summary

The inductance of an iron-cored coil to alternating current is generally very much smaller when the alternating current is superposed on a direct current. This has been shown to be due to the fact that the permeability of the iron is considerably reduced by the presence of direct current.

In order to obtain a larger inductance an air gap is introduced into the iron circuit. A method of calculating the required optimum air gap is worked out.

References.

(1) C. R. Hanna, "Design of Reactances and Transformers which carry Direct Current, Journ. Amer. Inst. E. E., Feb. 1927, p. 155.
(2) A. H. Symonds, "Loop Permeability in Iron, and the Optimum Air Gap in an Iron Choke with D.C. Excitation," 'Experimental Wireless,' Sept. 1928, p. 485.
(3) G. F. Partridge, "The Inductance of Iron-cored Coils having an Air Gap," Phil. Mag., Oct. 1936, p. 665.

Appendix.

An alignment chart has been constructed to determine the required length of gap l_a, knowing the area A, the length S, and the perimeter P of the iron core and the desired effective length of air gap l_e.

A is in sq. cm., S and P in cm., and l_e, l_a, etc., in mm.

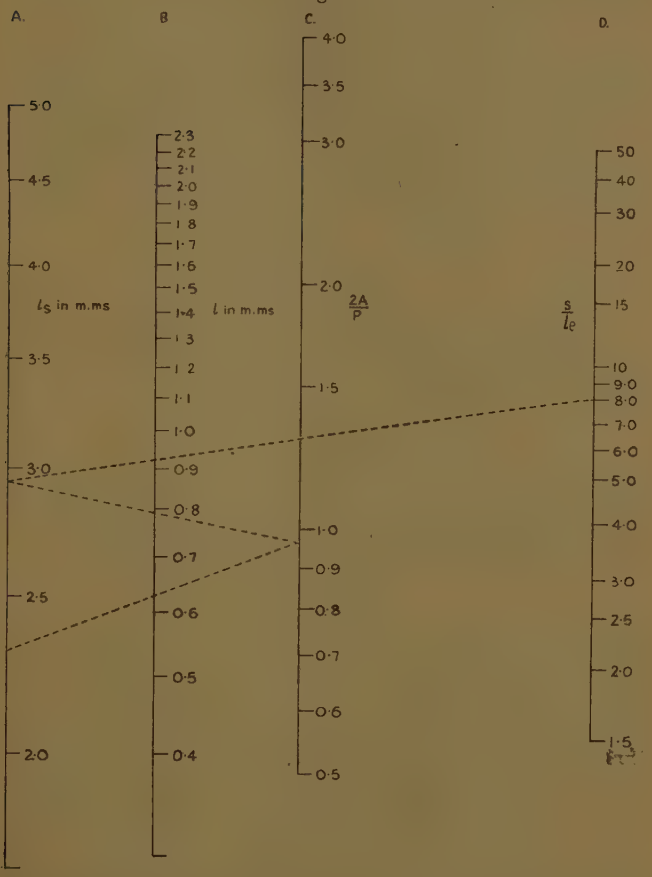
The equations which require to be solved are

$$l_s = \frac{2A}{P} \Big/ \log_e \frac{S}{l_e} \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

and $$l_a = \frac{l_e . l_s}{l_s - l_e} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

(*a*) To find l_s we have the three scales A, C, and D.

Fig. 3.



Having calculated the values of $\frac{S}{l_e}$ and $\frac{2A}{P}$ we select these values on the D and C scales respectively. A straight

line through these cuts the A scale at the required value of l_s.

(*b*) In order to evaluate l_a the A, B, and C scales are used as follows. Subtracting the value of l_e from the value of l_s already obtained as above, we get a lower point on the A scale corresponding to $l_s - l_e$. We now take a cut through the given value of l_e on the B scale to a point on the C scale used as a reference line only, and back again from this point to the value of l_s already obtained on the A scale.

The point at which this second line cuts the B scale is the required value of l_a.

As illustrating the method the following case has been shown on the chart :—

$S = 5{\cdot}0$ cm. $l_e = 0{\cdot}625$ mm. $\frac{S}{l} = 8{\cdot}0$,

$A = 6{\cdot}7$ sq. cm. $P = 10{\cdot}4$ cm. $\frac{2A}{P} = 1{\cdot}29$,

giving $l_s = 2{\cdot}94$ mm. as shown in (*a*) above.

Hence $l_s - l_e = 2{\cdot}31$ mm.,

giving $l_a = 0{\cdot}79$ mm. as shown in (*b*) above.

X. *On Waves in Canals of Variable Depth.* *By* B. R. Seth, *M.A., M.Sc., Ph.D.**

1. *Introduction.*

THE number of cases of transverse oscillations of water contained in canals of variable depth, of which the solution has been obtained, is very small. In fact, the only case whose complete solution is known is due to Kirchhoff †. It is when the section of the canal consists of two straight lines inclined at 45° to the vertical. When they are inclined at an angle of 60° to the vertical a symmetrical mode of oscillation due to Greenhill is known ‡. Rayleigh's method may be used to get an approximate value of the frequency of the slowest mode when the free surface is at the level of the axis of a circular

* Communicated by the Author.

† *Ges. Abh.* p. 428.

‡ See Lamb's 'Hydrodynamics.' pp. 419–420 (1930, 5th edition).

canal *. The object of the present paper is to show that the frequency of the gravest mode is $(g/h)^{\frac{1}{2}}$ in each of the following cases :—

(i.) When the section is a pair of straight lines equally inclined to the horizontal at an angle θ given by

$$\theta = \frac{\pi}{2p}, \frac{3\pi}{2p}, \frac{5\pi}{2p}, \ldots \frac{(p-r)\pi}{2p},$$

p being any integer, r being 1 or 2 according as p is even or odd, and h being the height of water in the canal measured from their point of intersection.

(ii.) When the section is a vertical straight line and a line inclined to the horizontal at an angle θ given by

$$\theta = \frac{\pi}{2p}, \frac{3\pi}{2p}, \frac{5\pi}{2p}, \ldots \frac{(p-2)\pi}{2p},$$

p being any odd integer.

(iii.) When the section is (a) a rectangular hyperbola, (b) a hyperbola whose eccentricity is 2, h being now measured from the centre of the hyperbolas.

2. *Conditions to be satisfied by the Velocity Potential.*

If we take the axes of x and y respectively horizontal and vertical in the plane of a cross-section, the problem of transverse oscillations reduces to the determination of a velocity potential, ϕ, satisfying the condition

$$\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} = 0 \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

throughout the liquid,

$$\frac{\partial\phi}{\partial n} = 0 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

at fixed boundaries, and

$$\sigma^2\phi = g\frac{\partial\phi}{\partial y} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

at the free surface, where δn is an element of the normal drawn to the fixed boundaries, and σ stands for the frequency of any mode.

* See Lamb's 'Hydrodynamics,' pp. 419–420 (1930, 5th edition).

3. *Rectilinear Section.*

Let the section consist of a pair of straight lines, each inclined at an angle π/m to the horizontal, m being an integer. If ψ is the stream function, we may write

$$\phi+i\psi = \mathrm{F}(z^m)\cos(\sigma t+\epsilon). \quad . \quad . \quad . \quad . \quad (4)$$

In (4) $\mathrm{F}(z^m)$ is a function of z^m which remains finite, one-valued, and continuous throughout the section. Hence it can be expanded in a power-series in ascending powers of z^m. From (2) we see that ψ must be zero over both $\theta=\pi/m$ and $\theta=\pi-\pi/m$. This condition is obviously satisfied, since from (4), omitting the time-factor $\cos(\sigma t+\epsilon)$, we have

$$\psi_1 = \mathrm{A}_1 r^{m/2}\cos\tfrac{1}{2}m\theta+\mathrm{A}_3 r^{3m/2}\cos\tfrac{3}{2}m\theta+\mathrm{A}_5 r^{5m/2}\cos\tfrac{5}{2}m\theta+\ldots, \quad . \quad . \quad . \quad (5.1)$$

$$\psi_2 = \mathrm{B}_1 r^m \sin m\theta+\mathrm{B}_2 r^{2m}\sin 2m\theta+\mathrm{B}_3\sin 3m\theta+\ldots, \quad . \quad (5.2)$$

with the corresponding values of ϕ given by

$$\phi_1 = \mathrm{A}_0-\mathrm{A}_1 r^{m/2}\sin\tfrac{1}{2}m\theta-\mathrm{A}_3 r^{3m/2}\sin\tfrac{3}{2}m\theta -\mathrm{A}_5 r^{5m/2}\sin\tfrac{5}{2}m\theta-\ldots, \quad (6.1)$$

$$\phi_2 = \mathrm{B}_0+\mathrm{B}_1 r^m\cos m\theta+\mathrm{B}_2 r^{2m}\cos 2m\theta +\mathrm{B}_3\cos 3m\theta+\ldots, \quad . \quad (6.2)$$

A's and B's being all real constants which are to be determined from the condition to be satisfied at the free surface given by (3).

If $m=2p$, and p be even, ϕ_2 gives the symmetrical oscillations and ϕ_1 the asymmetrical. If p be odd, ϕ_1 and ϕ_2 both give symmetrical oscillations. This is due to the vanishing of ψ_1, for $\theta=\frac{1}{2}\pi$ as well. If, therefore, we replace one of the inclined boundaries by the axis of y which is vertical, ϕ_1 gives the asymmetrical type for this section. It also follows that when p is odd the gravest mode for a canal whose section consists of two straight lines equally inclined to the horizontal will be symmetrical, and not asymmetrical. Lamb *, therefore, does not appear to be right when he says that in the case treated by Greenhill, in which the value of p is 3, the slowest mode must be asymmetrical.

* *Loc. cit.*, 'Hydrodynamics,' p. 419.

When m is odd, the series $F(z^m)$ in ascending powers of $z^{\frac{1}{m}}$ ceases to be one-valued, and hence ϕ_1 should be rejected. But now ϕ_2 itself splits up into two series

$$B_0+B_2r^{2m}\cos 2m\theta+B_4r^{4m}\cos 4m\theta+\ldots \quad . \quad (7.1)$$

and

$$B_1r^m\cos m\theta+B_3r^{3m}\cos 3m\theta+\ldots \quad . \quad . \quad . \quad (7.2)$$

which respectively give the symmetrical and the asymmetrical modes.

4. *Case of $m=2p$.*

Let h denote the height of the free surface above the origin, then, putting $\sigma^2/g=l$, (3) gives, if we use ϕ_1,

$$l[A_0-A_1r^p\sin p\theta-A_3r^{3p}\sin 3p\theta-A_5r^{5p}\sin 5p\theta-\ldots]$$
$$=-p[A_1r^{p-1}\cos(p-1)\theta+3A_3r^{3p-1}\cos(3p-1)\theta$$
$$+5A_5r^{5p-1}\cos(5p-1)\theta+\ldots]. \quad (8)$$

Each term on both sides of (8) can be expanded in rational, integral powers of x and y. Putting $y=h$, and equating the coefficients of different powers of x, we get, in general, an infinite number of equations to determine A's and l. The problem of determining σ, therefore, reduces to the evaluation of an infinite determinant.

(i.) *When p is an even integer—*

In this case $A_0=0$, and the coefficients of n, x^3, etc., respectively equated to zero give

$$\sum_{r=0}^{\infty}[(-1)^{\frac{1}{2}np}A_nh^{np-1}npC_1(lh-\overline{np-1})]=0, \quad . \quad (9.1)$$

$$\sum_{r=0}^{\infty}[(-1)^{\frac{1}{2}np}A_nh^{np-1}npC_3(lh-\overline{np-3})]=0, \quad . \quad (9.2)$$

$$\sum_{r=0}^{\infty}[(-1)^{\frac{1}{2}np}A_nh^{np-1}npC_5(lh-\overline{np-5})]=0, \quad . \quad (9.3)$$

where $n=2r+1$, it being understood that after $\frac{1}{2}p$ equations the infinite series begin with A_3, after $\frac{3}{2}p$ equations they begin with A_5, and so on.

Eliminating A's from (9) we get an infinite determinant to determine lh of the form

$$\begin{vmatrix} pc_1(lh-\overline{p-1}) & 3pc_1(lh-\overline{3p-1}) & 5pc_1(lh-\overline{5p-1}) & \dots \\ pc_3(lh-\overline{p-3}) & 3pc_3(lh-\overline{3p-3}) & 5pc_3(lh-\overline{5p-3}) & \dots \\ pc_5(lh-\overline{p-5}) & 3pc_5(lh-\overline{3p-5}) & 5pc_5(lh-\overline{5p-5}) & \dots \\ \dots & \dots & \dots & \dots \\ pc_{p-1}(lh-1) & 3pc_{p-1}(lh-\overline{2p-1}) & 5pc_{p-1}(lh-\overline{4p-1}) & \dots \\ 0 & 3pc_{p+1}(lh-\overline{2p-1}) & 5pc_{p+1}(lh-\overline{4p-1}) & \dots \\ \dots & \dots & \dots & \dots \\ 0 & 3pc_{2p+1}(lh-\overline{p-1}) & 5pc_{2p+1}(lh-\overline{3p-1}) & \dots \\ \dots & \dots & \dots & \dots \end{vmatrix} = 0, \quad \dots \quad (10.1)$$

which may be re-written as

$$\begin{vmatrix} \frac{1}{(p-1)!}(lh-\overline{p-1}) & \frac{1}{(3p-1)!}(lh-\overline{3p-1}) & \frac{1}{(5p-1)!}(lh-\overline{5p-1}) \dots \\ \frac{1}{(p-3)!}(lh-p-3) & \frac{1}{(3p-3)!}(lh-\overline{3p-3}) & \frac{1}{(5p-3)!}(lh-\overline{5p-3}) \dots \\ \frac{1}{(p-5)!}(lh-\overline{p-5}) & \frac{1}{(3p-5)!}(lh-\overline{3p-5}) & \frac{1}{(5p-5)!}(lh-\overline{5p-5}) \dots \\ \dots & \dots & \dots \\ 0 & \frac{1}{(p-1)!}(lh-p-1) & \frac{1}{(3p-1)!}(lh-\overline{3p-1}) \dots \\ 0 & \frac{1}{(p-3)!}(lh-p-3) & \frac{1}{(3p-5)!}(lh-3p-3) \dots \\ \dots & \dots & \dots \end{vmatrix} = 0 \quad \dots \quad (10.2)$$

In (10.2), after the first p rows the next p are obtained by displacing the first p through one column, the next p by displacing them through two columns, and so on. In fact, the determinant may be called a multi-gradient, and can be interpreted as the condition that the following p equations in t may have a root in common * :—

$$\sum_{r=0}^{\infty} \frac{t^{rp-1}}{(rp-1)!}(lh-\overline{rp-1}) = 0, \quad \dots \quad (11.1)$$

* This t should not be confused with the t in the time factor.

$$\sum_{r=0}^{\infty} \frac{t^{np-3}}{(np-3)!}(lh-\overline{np-3})=0, \quad \ldots\ldots\ldots \quad (11.2)$$

$$\sum_{r=0}^{\infty} \frac{t^{np-p+1}}{(np-p+1)!}(lh-\overline{np-p+1})=0, \quad (n=2r+1). \quad (11.3)$$

Each of these series can be summed up with the help of Simpson's method. If we notice that

$$\sum_{r=0}^{\infty} \frac{t^{2r+1}}{(2r+1)!}(lh-\overline{2r+1})=lh \sinh t-t \cosh t,$$

we easily see that (11) reduce to the following p independent equations :

$$\left.\begin{aligned} &lh \sinh t-t \cosh t=0,\\ &lh \sinh t\sqrt{\omega}-t\sqrt{\omega} \cosh t\sqrt{\omega}=0,\\ &lh \sinh t\sqrt{\omega^2}-t\sqrt{\omega^2} \cosh t\sqrt{\omega^2}=0,\\ &\quad \ldots\ldots\ldots\ldots\ldots\\ &lh \sinh t\sqrt{\omega^{p-1}}-t\sqrt{\omega^{p-1}} \cosh t\sqrt{\omega^{p-1}}=0, \end{aligned}\right\} \quad (12)$$

where ω is any pth root of unity excepting 1. The value of $\phi+i\psi$ is now given by *

$$\phi+i\psi=iA[(\omega^0)^{\frac{1}{p}} \cos kz+(\omega^1)^{\frac{1}{p}} \cos kz\sqrt{\omega} + (\omega^2)^{\frac{1}{p}} \cos kz\sqrt{\omega^2}+\ldots(\omega^{p-1})^{\frac{1}{p}} \cos kz\sqrt{\omega^{p-1}}], \quad (13)$$

omitting the time factor $\cos(\sigma t+\epsilon)$. In (13) A and k are both real constants.

Before proceeding further we may verify the above results by the known case of $p=2$. (12) and (13) now reduce to

$$lh \sinh t-t \cosh t=0, \quad lh \sin t-t \cos t=0,$$

and

$$\phi+i\psi=iA[\cos kz-\cosh kz], \quad (t=kh).$$

These results are well known.

The case of $p=1$ is trifling. Barring these two values of p we see that the p equations in (12) cannot all be satisfied by any real value of t excepting zero. $t=0$ gives $lh=1$, and hence $\sigma=(g/h)^{\frac{1}{2}}$. Apparently no other type of mode is stable.

* In general, we shall have to replace k in this value of $\phi+i\psi$ by $k\alpha_1$, $k\alpha_2$, etc., add all these solutions, and choose α's such that when $k\to 0$ the resultant ϕ has all finite real coefficients and satisfies (3).

(ii.) *When p is odd—*

Proceeding as in (i.) we get the same p equations as in (12), but the value of $\phi+i\psi$ now takes the form *

$$\phi+i\psi = iA[(\omega^0)^{\frac{1}{2}p}\sin kz+(\omega^1)^{\frac{1}{2}p}\sin kz\sqrt{\omega} +(\omega^2)^{\frac{1}{2}p}\sin kz\sqrt{\omega^2}+\ldots(\omega^{p-1})^{\frac{1}{2}p}\sin kz\sqrt{\omega^{p-1}}]. \quad (14)$$

Thus, in this case, too, the only type of mode that is stable is given by $\sigma=(g/h)^{\frac{1}{2}}$.

5. *Types of Mode when ϕ_2 is used.*

In Section 4 we have used ϕ_1 given by (6.1). But if we use ϕ_2 given by (6.2), our equations similar to (12) come out to be

$$\left.\begin{aligned} &lh\cosh t - t\sinh t = 0,\\ &lh\cosh t\sqrt{\omega} - t\sqrt{\omega}\sinh t\sqrt{\omega} = 0,\\ &lh\cosh t\sqrt{\omega^2} - t\sqrt{\omega^2}\sinh t\sqrt{\omega^2} = 0,\\ &\quad\ldots\ldots\ldots\ldots\\ &lh\cosh t\sqrt{\omega^{p-1}} - t\sqrt{\omega^{p-1}}\sinh t\sqrt{\omega^{p-1}} = 0, \end{aligned}\right\} \quad (15)$$

and the value of $\phi+i\psi$ becomes *

$$\phi+i\psi = A[\cos kz+\cos kz\sqrt{\omega}+\cos kz\sqrt{\omega^2}+\ldots. \cos kz\sqrt{\omega^{p-1}}], \quad (16.1)$$

(p even),

$$\phi+i\psi = A(\sin kz+\sin kz\sqrt{\omega}+\sin kz\sqrt{\omega^2}+\ldots \sin kz\sqrt{\omega^{p-1}}] \quad (16.2)$$

(p odd).

For $p=2$, (15) and (16) reduce to Kirchhoff's results. In all other cases the only common real root of the p equations in (15) is $t=0$, which gives $\sigma=0$. This only amounts to an alteration of the origin. Barring the case of $p=2$, we therefore get no stable mode for any other case.

6. *When m is an odd Integer.*

In this case, too, we must use the two forms into which ϕ_2 splits up, and, proceeding as above, we find that no stable mode can be obtained.

* In general, we shall have to replace k in this value of $\phi+i\psi$ by $k\alpha_1$, $k\alpha_2$, etc., add all these solutions, and choose α's such that when $k\to 0$ the resultant ϕ has all finite real coefficients and satisfies (3).

7. *Canal bounded by a Vertical Wall.*

As we have already pointed out in Section 3, the value of ϕ_1 given by (6.1) may also be used for a canal whose section is a vertical straight line and another inclined at an angle θ given by

$$\theta = \frac{\pi}{2p}, \frac{3\pi}{2p}, \frac{5\pi}{2p}, \ldots \frac{(p-2)\pi}{2p},$$

provided we assume that p is an odd integer. We have already treated this case in Section 4, where we have found that in all cases the only stable mode is given by

$$\sigma = (g/h)^{\frac{1}{2}}.$$

8. *Hyperbolic Section.*

(*a*) If the section is a rectangular hyperbola given by

$$y^2 - x^2 = a^2,$$

we can take

$$\phi = 2\mathrm{A}xy \cos(\sigma t + \epsilon), \quad \psi = \mathrm{A}(y^2 - x^2 - a^2) \cos(\sigma t + \epsilon) \quad . \quad (17)$$

The surface condition (3) now gives the asymmetrical mode

$$\sigma = (g/h)^{\frac{1}{2}},$$

h being measured from the centre of the hyperbola.

The corresponding form of the free surface is

$$\zeta = \frac{1}{g}\left[\frac{\partial \phi}{\partial t}\right]_{y=h} = -\frac{2\mathrm{A}h\sigma}{g} x \sin(\sigma t + \epsilon).$$

The surface in this mode is therefore always plane. If $a = 0$, we get the slowest mode of Kirchhoff's solution.

(*b*) If the section is a hyperbola given by

$$y^2 - \tfrac{1}{3}x^2 = a^2,$$

we can take

$$\left.\begin{aligned} \phi &= [\mathrm{A}_0 - \mathrm{A}y(\tfrac{1}{3}y^2 - x^2 - a^2)] \cos(\sigma t + \epsilon), \\ \psi &= \mathrm{A}x(y^2 - \tfrac{1}{3}x^2 - a^2) \cos(\sigma t + \epsilon). \end{aligned}\right\} \quad . \quad (18)$$

The condition at the free surface gives the symmetrical mode

$$\sigma = (g/h)^{\frac{1}{2}}, \quad \mathrm{A}_0 = -\tfrac{2}{3}\mathrm{A}h^3.$$

The corresponding form of the free surface

$$\zeta = \frac{\mathrm{A}h\sigma}{g}(h^2 - x^2 - a^2) \sin(\sigma t + \epsilon)$$

is a parabolic cylinder with two nodes at distances of 0·5774 of the half-breadth from the centre. If we put $a=0$, the case degenerates into Greenhill's solution.

Since ψ given in (18) vanishes for $x=0$, we can also use this solution when the section is bounded by a vertical straight line and the arc of the hyperbola lying in a quadrant. In this case the mode becomes asymmetrical.

Hindu College,
Delhi, India.

XI. *New Basic Critical Equations and Methods in the Calculus of Variations.* *By* W. S. Kimball, *Ph.D.*, *Michigan State College, East Lansing, Michigan* *.

Introduction.

NEW conditions are established by a delta process of four stages similar to that employed in elementary calculus. In the first stage we "set up" an integral approximately as a sum in terms of n pairs of interval variables Δx, Δy. In the next stage the Lagrange multiplier method for conditional extremes is applied to the "set up" *before* passing to the limit. The critical equations of the Lagrange method as applied to the finite sum become in general as we pass to the limit the new critical differential equations for an extremum of the now integral limit :

$$\left.\begin{aligned} \frac{\partial J}{\partial dx} &= \frac{\partial F}{\partial dx} - \frac{\partial G}{\partial dx} = \frac{\partial F}{\partial dx} - c_1 \\ &= \int_x^{x_n} f_x\,dx + f - y'f_{y'} - c_1 = 0, \\ \frac{\partial J}{\partial dy} &= \frac{\partial F}{\partial dy} - \frac{\partial G}{\partial dy} = \frac{\partial F}{\partial dy} - c_2 \\ &= \int_x^{x_n} f_y\,dx + f_{y'} - c_2 = 0. \end{aligned}\right\} \quad . \quad . \quad . \quad (48)$$

These critical equations exhibit new features in that (a) they are special first-order equations representing

* Communicated by the Author.

differentiation with respect to differentials. They include * the Du Bois Raymond equation (right members of the second of (48)) and reduce upon differentiation to Euler's equations including also the factors y' and $1/y'$. (*b*) They combine the older method of differentials with the Lagrange multiplier method, using an isoperimetric vector condition on the differentials

$$G=c_1[\int dx-(x_n-x_0)]+c_2[\int dy-(y_n-y_0)]=0.$$

(*c*) The method does not use or rely on the fundamental lemma of the calculus of variations, nor the " variations " of Euler and Lagrange, such as δf and $\delta y=\eta(x)$ †. (*d*) In physics these conditions represent conservation of energy and momentum according to Hamilton's Principle, and upon differentiation yield the force equations of mechanics. (*e*) The whole spirit and method consists, as indicated by the left members of (48), in differentiation with respect to differentials, for which in connexion with the Lagrange method, new working rules (I.) and (II.) of Section 10 are established and illustrated. (*f*) A mathematical test is the sole test for admissible arcs and functions, which include any that provide existing test formulas (48), (49), and (53) and the needful solutions.

It is conjectured ‡ that (48) are the true inclusive necessary conditions for an extremum of an integral (49) because all known maximizing and minimizing arcs are found to be solutions without exception; and this includes all the special cases that are not solutions of Euler's equations, the latter being hence not " necessary " conditions in the strict and rigorous sense of the word. Sufficiency conditions are not considered.

1. *The Multiplier Method for Conditional Maxima and Minima.*

We begin by recalling the essential features of the Lagrange multiplier method. The problem of maximizing a function $F=xyz$, subject to a constant condition

* See Bolza, ' Variationsrechnung,' p. 30.

† Whose commutativity with d and $\int$ " forms the basis " of the calculus of variations according to A. G. Webster, ' Dynamics,' p. 80. See also Bolza, *loc. cit.* §8, p. 45.

‡ Without proof which must include every unimagined possibility in the way of minimizing and maximizing arcs.

$G=\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}-1=0$, involves only two independent variables, say x and y, because of the restriction $G=0$. Hence the critical equations for the maximum of F are

$$\left.\begin{aligned}\frac{\partial F}{\partial x}&=F_x+F_z\frac{\partial z}{\partial x}=0,\\ \frac{\partial F}{\partial y}&=F_y+F_z\frac{\partial z}{\partial y}=0,\end{aligned}\right\} \quad \ldots\ldots \quad (1)$$

where F_x, F_y, and F_z refer to the partial derivatives of F with respect to x, y, and z treated as independent variables. Besides (1) we have from the condition $G=0$:

$$\left.\begin{aligned}\frac{\partial G}{\partial x}&=G_x+G_z\frac{\partial z}{\partial x}=0,\\ \frac{\partial G}{\partial y}&=G_y+G_z\frac{\partial z}{\partial y}=0.\end{aligned}\right\} \quad \ldots\ldots \quad (2)$$

With the help of (2) we may eliminate $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ from (1) and obtain

$$\left.\begin{vmatrix}F_x & G_x\\ P_z & G_z\end{vmatrix}=0, \quad \begin{vmatrix}F_y & G_y\\ F_z & G_z\end{vmatrix}=0,\right\} \quad \ldots \quad (3)$$

which for the particular F and G here chosen yield

$$\frac{x^2}{a^2}=\frac{y^2}{b^2}=\frac{z^2}{c^2}$$

for the critical values of x, y, and z that maximize F subject to $G=0$.

Now a curious algebraic anomaly appears when we obtain the *unrestricted* maximum of $J(xyz)=F+\lambda G$ (where λ is an undetermined constant). The critical equations are given by the vanishing of the partials with respect to the now three independent variables

$$\frac{\partial J}{\partial x}=F_x+\lambda G_x=0, \qquad \frac{\partial J}{\partial y}=F+\lambda G_y=0,$$

$$\frac{\partial J}{\partial z}=F_z+\lambda G_z=0. \quad \ldots \quad (4)$$

The elimination of λ from (4) gives nothing more or less than (3), so that these critical equations for the problem of maximizing F subject to a constant G are *identical* with those that maximize the unrestricted $J=F+\lambda G$ according to (4), and hence the two problems are equivalent.

It will readily be seen that the above procedure works equally well with any number of variables. Accordingly the maximum of $F(x_1, x_2 \ldots x_n)$ subject to a constant condition $G(x_1, x_2 \ldots x_n)$ is determined by $n-1$ first-order determinants :

$$\left.\begin{vmatrix} F_{x_1} & G_{x_1} \\ F_{x_n} & G_{x_n} \end{vmatrix} = 0, \quad \begin{vmatrix} F_{x_2} & G_{x_2} \\ F_{x_n} & G_{x_n} \end{vmatrix} = 0, \ldots, \begin{vmatrix} F_{x_{n-1}} & G_{x_{n-1}} \\ F_{x_n} & G_{x_n} \end{vmatrix} = 0, \right\} \quad (5)$$

which are *identically* the very same conditions obtained by eliminating λ from the n equations that represent the vanishing of the partials of $J=F+\lambda G$ with respect to the n *independent* variables

$$F_{x_1}+\lambda G_{x_1}=0 : \ F_{x_2}+\lambda G_{x_2}=0 ; \ \ldots, F_{x_n}+\lambda G_{x_n}=0. \ . \quad (6)$$

Hence equations (6) represent the maximum of F subject to a constant $G=0$, where λ is a constant to be determined like a constant of integration by the particular conditions that we choose to impose on our solution, being tied up with $G=0$.

2. *After an integral is set up (Stage I.) in terms of n pairs of finite interval variables Δx and Δy, the multiplier method (Stage II.) may be applied* before *passing to the limit* *.

Since problems in the Calculus of Variations generally involve the maximum or minimum of a more or less *definite integral*, they are *ipso facto* conditional, and of the sort to which the multiplier method applies.

We may illustrate by a determination of the shortest distance between the two points 1 and 2. We attach labels I., II., and III. and IV. to the four stages to be emphasized in the procedure.

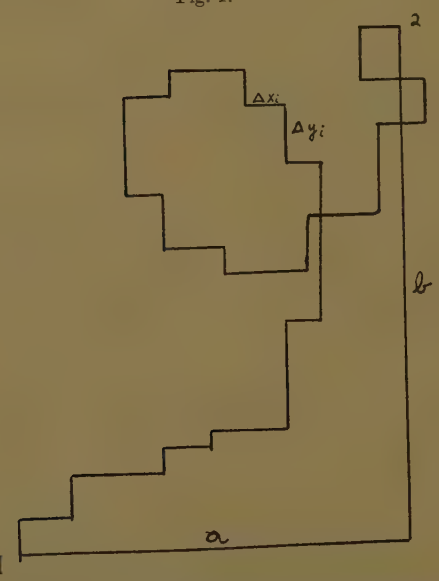
* This method is the same as the conditional maximum problem of the geometrical weight method (Kimball, W. S., Journ. Phys. Chem. xxxiii. p. 1558 (1929)) in that a differential equation satisfied by n critical constants is integrated to get the solution.

Summation Stage I.—Let the space be broken up into pairs of horizontal and vertical interval variables Δx_i and Δy_i with i ranging from 1 to n, as indicated in fig. 1. Let the *number* of variables be fixed and let us minimize the sum

$$\mathrm{F} = \sum_{i=1}^{n} \sqrt{\Delta x_i^2 + \Delta y_i^2},$$

$$\text{subject to } \mathrm{G} \equiv c_1 \left[\sum_{1}^{n} \Delta x_i - a \right] + c_2 \left[\sum_{1}^{n} \Delta y_i - b \right] = 0. \quad (7)$$

Fig. 1.



Interval variables and length elements that comprise F.

This G=0 is a constant vector restriction, equivalent to the two scalar restrictions $\sum_{i=1}^{n} \Delta x_i = a$ and $\sum_{i=1}^{n} \Delta y_i = b$. It does nothing more than specify that F or its limit (12) shall be a definite sum or integral—definite as to its y interval as well as its x interval. Note that analysis

involving path variations between these limits is essentially *vector analysis*, since the *direction* of the paths is the crux of the problem.

The two constants in G=0 may be written

$$c_1=\lambda_1 i \quad \text{and} \quad c_2=\lambda_2 j, \quad . \quad . \quad . \quad . \quad . \quad (8)$$

where i and j are the familiar constant unit vectors in the x and y directions and the λ's are the undetermined multipliers of the Lagrange method. These λ's are incorporated in G=0 instead of appearing explicitly as in (4) and (6), and also we take J as the *difference* F–G because this usually leads to positive c's and greater analytical simplicity.

Arithmetic Stage II.—First keep all the Δy's depicted in fig. 1 *constant* and treat F as a function of the n variable Δx's. Apply the Lagrange multiplier method and set equal to zero the n partial derivatives of J=F−G given by (7) with respect to the $n\,\Delta x_i$, taking note that $\frac{\partial G}{\partial \Delta x^i}=c_1$.

$$\frac{\partial J}{\partial \Delta x_i}=\frac{\Delta x_i}{\sqrt{\Delta x_i^2+\Delta y_i^2}}-c_1=0. \quad . \quad . \quad . \quad (9)$$

This set of n eqs. (9) together with G=0, suffices to determine each Δx_i as well as c_1 as critical "constants" in terms of the angle of inclination α and the "constants" Δy_i. We find

$$c_1=\frac{a}{\sqrt{a^2+b^2}}=\cos\alpha\,; \quad \Delta x_i=\Delta y_i \cot\alpha, \quad . \quad (10)$$

indicating how each "critical constant" Δx_i must be related to its arbitrary Δy_i (plus or minus as in fig. 1) so as to minimize (7), *i. e.*, make it represent a straight line.

Algebraic Stage III.—Drop the subscripts from (9),

$$\frac{\partial J}{\partial \Delta x}=\frac{\Delta x}{\sqrt{\Delta x^2+\Delta y^2}}-c_1=0. \quad . \quad . \quad . \quad (11)$$

Here the Δx and Δy appear as the *variables* in an algebraic equation that is satisfied by every pair of constants Δx_i and Δy_i determined by (9) as well as infinitely many more such pairs. Note the contrast between (9) and (11), the

former being n eqs. that determine n pairs of constants each pair of which will satisfy the single algebraic (11) with its one pair of variables, and that the *transition* from (9) to (11) is like the transition from arithmetic to algebra. Hence the headings attached to these two stages.

Calculus Stage IV.—Pass to the limit as Δx approaches zero. Then (7) and (11) become the integrals (12) and the differential equation (13).

$$F=\int_1^2 \sqrt{dx^2+dy^2}\,;\quad G\equiv c_1[\int dx-a]+c_2[\int dy-b]=0. \quad (12)$$

$$\frac{\partial J}{\partial dx}=\frac{\partial F}{\partial dx}-\frac{\partial G}{\partial dx}=\frac{dx}{\sqrt{dx^2+dy^2}}-c_1=0. \quad . \quad (13)$$

Here (13) is the critical condition that determines a minimum of F subject to G according to (12), having the straight line solution indicated by (14)

$$y=x\tan\alpha. \quad . \quad . \quad . \quad . \quad . \quad . \quad (14)$$

Likewise we might have kept the Δx_i's constant instead of the Δy_i's as in (9) and obtained the partial derivatives with respect to Δy_i and dy

$$\frac{\partial J}{\partial dy}=\frac{dy}{\sqrt{dx^2+dy^2}}-c_2=0\,;\quad c_2=\sin\alpha. \quad . \quad (15)$$

Eqns. (13) and (15) are the new critical conditions for an extremum of F subject to the constant condition $G=0$.

3. *Shortest Distance Problems in Polar Coordinates.*

Summation Stage I.—As with cartesian coordinates we set up a finite sum in terms of a fixed number of finite interval variables which are subjected to a constant vector restriction * in place of the definite integral. Thus

$$F=\sum_{i=1}^{n}\sqrt{\Delta r_i^2+r_i^2\Delta\theta_i^2}\,;$$

$$G=c_1[\Sigma\Delta\theta_i-(\theta_n-\theta_0)]+c_2[\Sigma\Delta r_i-(b-a)]=0. \quad (16)$$

* The equivalent of two separate scalar conditions, one for each set of variables.

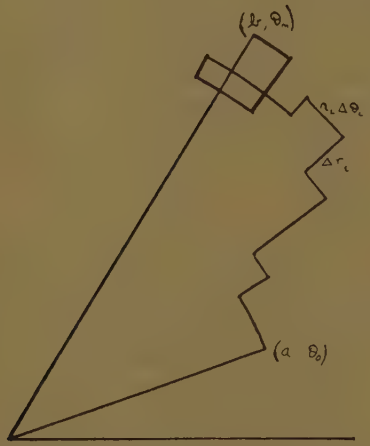
Arithmetic Stage II.—Apply the Lagrange method by setting equal to zero the partials of J=F−G with respect to the n independent $\Delta\theta_i$

$$\frac{\partial J}{\partial \Delta\theta^i} = \frac{r_i^2 \Delta\theta_i}{\sqrt{\Delta r_i^2 + r_i^2 \Delta\theta_i^2}} - c_1 = 0. \qquad \frac{\partial G}{\partial \Delta\theta_i} = c_1. \quad . \quad (17)$$

Eqs. (17), like (9), are n equations that determine the n critical constant values of the variables $\Delta\theta_i$ in terms of the n other arbitrary fixed " constants " Δr_i so as to make an extremum of (16) where

$$r_i = a + \sum_{j=1}^{i} \Delta r_j, \quad . \quad . \quad . \quad . \quad . \quad (18)$$

Fig. 2.



Polar interval variables and approximate length elements of F according to (16).

showing Δr_i as the last interval variable added to r_i as indicated by fig. 2.

Algebraic Stage III.—Now drop the subscripts from (17), as done with (11), and interpret Δr and $\Delta\theta$ as finite *variables* instead of constants, then instead of (17) we have the algebraic equation in Δr and $\Delta\theta$

$$\frac{\partial J}{\partial \Delta\theta} = \frac{r^2 \Delta\theta}{\sqrt{\Delta r^2 + r^2 \Delta\theta^2}} - c_1 = 0. \quad . \quad . \quad . \quad (19)$$

As with (11) the n pairs of constants Δr_i and $\Delta\theta_i$ determined by (17) are solutions of (19), which also, like (11), has infinitely many others. The left members of (9) and (17) are partial derivatives. Likewise the left members of (11) and (19) represent exactly what their form indicates, derivatives with respect to Δx and $\Delta\theta$, variables which can be any one of the Δx_i or $\Delta\theta_i$ of the former equations as well as other variables included by (7), (16), and (18) as n the number of variables ranges through all positive integers. Thus (11) and (19) are a generalization of (9) and (17), including the use of variables to represent the sets of constants of these latter, and the J, F, and G of (11) and (19) are functions of Δx and $\Delta\theta$ exactly as (7) and (16) are functions of Δx_i and $\Delta\theta_i$.

Calculus Stage IV.—Now pass to the limit as Δr and $\Delta\theta$ approach zero. Then (19) becomes the new critical *differential* equation (20) necessary for a minimum of the restricted or definite integral (21).

$$\frac{\partial J}{\partial d\theta} = \frac{r^2 d\theta}{\sqrt{dr^2+r^2 d\theta^2}} - c_1 = 0. \quad . \quad . \quad . \quad (20)$$

$$F = \int \sqrt{dr^2+r^2 d\theta^2}\,;$$

$$G = c_1[\int d\theta - (\theta_n - \theta_0)] + c_2[\int dr - (b-a)] = 0. \quad (21)$$

Eqs. (20), like (13) and (15) above and (24) below, represent what their form * indicates, *i. e.*, differentiation with respect to differentials, and in view of (18) we make no radical departure from the customary viewpoint that considers an integral as the limit of a sum, when we treat $J = F - G$ as a function *only* of the differentials that are involved.

Arithmetic Stage II.—Likewise we may take the partials of $F - G$ given by (16) with respect to Δr_i instead of $\Delta\theta_i$ according to stage II. :

$$\frac{\partial J}{\partial \Delta r_i} = \frac{\Delta r_i}{\sqrt{\Delta r_i^2 + r_i^2 \Delta\theta_i^2}} + \sum_{j=i}^{n} \frac{r_j \Delta\theta_j^2}{\sqrt{\Delta r_j^2 + r_j^2 \Delta\theta_j^2}} - c_2 = 0. \quad (22)$$

Note that the summation term results from (18), which shows by elementary calculus that $\frac{\partial r_j}{\partial \Delta r_i} = 0$ or 1 according

* See sections 9 and 10 below.

as j is less than i or otherwise, as indicated by fig. 2. Here again (22), like (17), is n equations that determine n critical pairs of constants Δr_i and $\Delta\theta_i$, the latter having been kept fixed and arbitrary, while (22) determines Δr_i in terms of them, instead of *vice versa* as done by (17).

Algebraic Stage III.—Now drop the subscripts from (22) and get a single algebraic equation in the variables Δr, $\Delta\theta$, and r :

$$\frac{\partial J}{\partial \Delta r}=\frac{\Delta r}{\sqrt{\Delta r^2+r^2\,\Delta\theta^2}}+\sum_{j=i}^{n}\frac{r\,\Delta\theta^2}{\sqrt{\Delta r^2+r^2\,\Delta\theta^2}}-c_2=0, \quad (23)$$

where the summation is to be taken from the particular pair Δr, $\Delta\theta$ of the first term (shown in fig. 2) up to the upper limit n, as indicated by the way (23) emerges from (22). Note also that we are now treating Δr and $\Delta\theta$ as variables whose number as well as magnitude may change subject only to (18) and G of (16), so that n can range through all positive integral values.

Calculus Stage IV.—Now pass to the limit and get (24) as an integral differential equation and necessary condition for the minimum of (21) :

$$\frac{\partial J}{\partial dr}=\frac{dr}{\sqrt{dr^2+r^2d\theta^2}}+\int_{\theta}^{\theta_n}\frac{r\,d\theta}{\sqrt{dr^2+r^2\,d\theta^2}}\,d\theta-c_2=0. \quad (24)$$

Eqs. (20) and (24) are the two new critical conditions for the minimum of the definite integral F of (21). The former is easily solved and yields

$$r=c_1\sec(\theta-c),$$

indicating c_1 as the perpendicular distance of this straight line from the origin when $\theta=c$. Eq. (24) is also easily solved by setting $\tan\psi=\frac{r\,d\theta}{dr}$, where ψ is the angle between radius vector and curve, then (24) becomes

$$\cos\psi+\int_{\theta}^{\theta_1}\sin\psi(\theta)\,d\theta-c_2=0,$$

which, upon differentiating with respect to θ, integrates to

$$\psi+\theta=\alpha$$

indicating a constant angle of inclination α for the hence straight line solution.

4. *Maximum Area enclosed by a given Arc Length.*

Consider the historical isoperimetric problem of maximizing the area that can be enclosed by a given arc length L.

I. We have to maximize

$$F=\sum_{i=1}^{n}\sum_{j=i}^{n}\Delta x_i\Delta y_j; \quad . \quad . \quad . \quad . \quad . \quad (25)$$

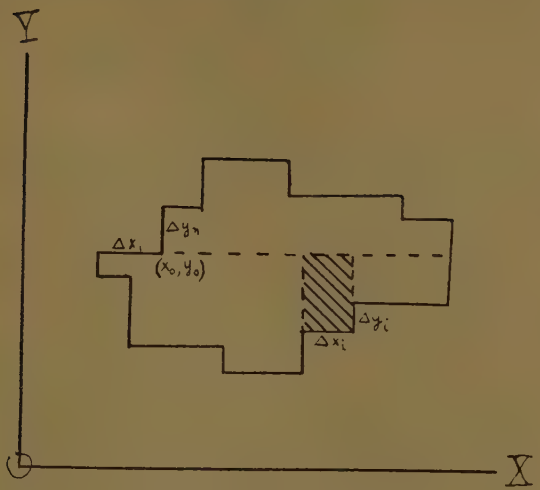
$$\left.\begin{aligned} &\text{subject to} \quad G\equiv c_1\sum_{i=1}^{n}\Delta x_i+c_2\sum_{i=1}^{n}\Delta y_i=0,\\ &\text{and} \quad H=\sum_{i=1}^{n}\sqrt{\Delta x_i^2+\Delta y_i^2}-L=0. \end{aligned}\right\} \quad . \quad . \quad . \quad (26)$$

F is seen to be the sum of such shaded areas indicated in fig. 3, and given by $\Delta F_i=\Delta x_i\sum_{j=i}^{n}\Delta y_j$. Like the vector constants (7) and (16) that specify the ranges of the two independent variables in the straight line problem, so here the vector $G=0$ is equivalent to two scalar equations $\Sigma\Delta x_i=0$; $\Sigma\Delta y_i=0$ which require our length L to be measured along a *closed* broken line or curve. Thus Δx_1, Δy_1 start at (x_0y_0) in fig. 3 where Δx_n, Δy_n terminate. Here also (8) applies whereby the Lagrange multipliers are included in the c's. Note also the isoperimetric scalar condition $H=0$.

II. The Lagrange method is known to apply equally well when more than one constant condition is involved. So, as before, we form the function $J=F-G+\lambda H$, where λ is the additional undetermined constant and differentiate with respect to the *n independent* variables Δx_i or Δy_i:

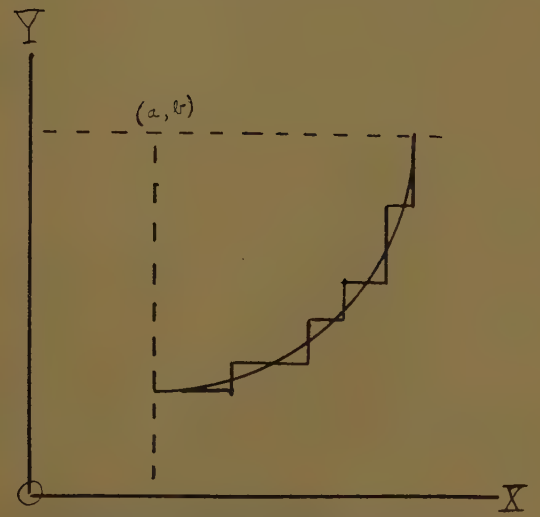
$$\left.\begin{aligned} \frac{\partial J}{\partial\Delta x_i}&=\frac{\partial F}{\partial\Delta x_i}-\frac{\partial G}{\partial\Delta x_i}+\lambda\frac{\partial H}{\partial\Delta x_i}\\ &=\sum_{j=i}^{n}\Delta y_j-c_1+\lambda\frac{\Delta x_i}{\sqrt{\Delta x_i^2+\Delta y_i^2}}=0,\\ \frac{\partial J}{\partial\Delta y_i}&=\frac{\partial F}{\partial\Delta y_i}-\frac{\partial G}{\partial\Delta y_i}+\lambda\frac{\partial H}{\partial\Delta y_i}\\ &=\sum_{j=1}^{i}\Delta x_j-c_2+\lambda\frac{\Delta y_i}{\sqrt{\Delta x_i^2+\Delta y_i^2}}=0. \end{aligned}\right\} \quad . \quad . \quad (27)$$

Fig. 3.



The broken line that encloses F.

Fig. 4.



The circular arc and broken line that maximize but do not enclose F when condition G=0 is omitted.

Note that the summation term ranges from $j=1$ to i in the second equation and in the first from i to n. This is because Δy_i is only present in those terms

$$\Delta \mathrm{F}_j = \Delta x_j \sum_{k=j}^{n} \Delta y_k,$$

for which j is less than or equal to i.

III. Now drop the subscripts and obtain *two* algebraic equations in the variables Δx and Δy :

$$\left.\begin{aligned} \frac{\partial \mathrm{J}}{\partial \Delta x} &= \sum_{j=i}^{n} \Delta y - c_1 + \lambda \frac{\Delta x}{\sqrt{\Delta x^2 + \Delta y^2}} = 0, \\ \frac{\partial \mathrm{J}}{\partial \Delta y} &= \sum_{j=1}^{i} \Delta x - c_2 + \lambda \frac{\Delta y}{\sqrt{\Delta x^2 + \Delta y^2}} = 0. \end{aligned}\right\} \quad . \quad . \quad (28)$$

As with (23) the first summation starts with the particular pair Δx, Δy of the right-hand term (not summed), and in the second equation (28) the summation ends with this pair. Each of these two equations will be satisfied by either set of n pairs Δx_i, Δy_i determined by the first or second of eqs. (27) as well as infinitely many more such sets, and J, F, and G are as in (11) and (19) functions of the sets of variables Δx, Δy included by (25) and (26) where n is any positive integer.

IV. When we pass to the limit we have, by elementary calculus, for the summation terms of (27) and (28), $\int_{y}^{y_0} dy = y_0 - y$ and $\int_{x_0}^{x} dx = x - x_0$, since (x_0, y_0) is the lower as well as upper limit of the sums (26) involved in $\mathrm{G}=0$.

Eqs. (28) now become

$$\left.\begin{aligned} \frac{\partial \mathrm{J}}{\partial dx} &= y_0 - y - c_1 + \frac{\lambda\, dx}{\sqrt{dx^2 + dy^2}} = 0, \\ \frac{\partial \mathrm{J}}{\partial dy} &= x - x_0 - c_2 + \frac{\lambda\, dy}{\sqrt{dx^2 + dy^2}} = 0, \end{aligned}\right\} \quad \left.\begin{aligned} \frac{dy}{dx} &= -\frac{x-a}{y-b} \\ &= -\frac{x-x_0-c_2}{y-y_0+c_1}, \end{aligned}\right\}$$

. . . (29)

where on the right we have set $a = a_0 + c_2$; $b = y_0 - c_1$. Simultaneously with (29) the sums (25) and (26) become the integral limits

$$\mathrm{F} = \iint dx\, dy \; ; \quad \mathrm{G} = \int_{x_0}^{x_0} (c_1 + c_2 y')\, dx = 0 \; ;$$

$$\mathrm{H} = \int \sqrt{1 + y'^2}\, dx - \mathrm{L} = 0. \quad . \quad (30)$$

Eqs. (29) are the new critical conditions for an extremum of the "definite" integral F subject according to (30) to the conditions G and H, the former of which merely specifies the limits of integration of F. The integration of each gives the circle of radius λ

$$(x-a^2)+(y-b)^2=\lambda^2, \quad . \quad . \quad . \quad . \quad (31)$$

which involves three arbitrary parameters.

It is of interest to note that if we omitted condition G=0 of (26) and used J=F+λH instead of as above, then (x_0y_0) would by (29) become equal to (a, b), the co-ordinates of the centre. Thus the end points of L given by (26) would be forcibly separated by the analysis and located on the centre lines $x=a$ and $y=b$ as indicated in fig. 4, whereby L would not be the length of a *closed* curve.

Likewise we may take F as the sum of the areas of triangles with vertices (o, o) $(x_i-\Delta x_i, y_i-\Delta y_i)$ and (x_i, y_i):

$$F=\frac{1}{2}\sum_{i=1}^{n}(x_i\Delta y_i-y_i\Delta x_i); \quad \text{and} \quad A=\tfrac{1}{2}\int(x\,dy-y\,dx), \quad (32)$$

using (26) again *, we have J=F−G+λH. Hence:

$$\left.\begin{aligned}\frac{\partial J}{\partial \Delta x_i}&=\frac{1}{2}\Big(\sum_{j=i}^{n}\Delta y_j-y_i\Big)-c_1+\frac{\lambda\Delta x_i}{\sqrt{\Delta x_i^2+\Delta y_i^2}}=0,\\ \frac{\partial J}{\partial \Delta y_i}&=\frac{1}{2}\Big(x_i-\sum_{j=i}^{n}\Delta x_j\Big)-c_2+\frac{\lambda\Delta y_i}{\sqrt{\Delta x_i^2+\Delta y_i^2}}=0,\end{aligned}\right\}. \quad (33)$$

which become in the limit, according to Stage IV.,

$$\left.\begin{aligned}\frac{\partial J}{\partial dx}&=\tfrac{1}{2}(y_0-2y)-c_1+\frac{\lambda\,dx}{\sqrt{dx^2+dy^2}}=0,\\ \frac{\partial J}{\partial dy}&=\tfrac{1}{2}(2x-x_0)-c_2+\frac{\lambda\,dy}{\sqrt{dx^2+dy^2}}=0,\end{aligned}\right\}. \quad (34)$$

which have the same solution (31) if we take $a=\frac{x_0}{2}+c_2$; $b=\frac{y_0}{2}-c_1$, indicating (34) as the necessary conditions for an extremum of (32) subject to (26).

* Note that in (33) the sums occur since by (45) below $\frac{\partial x_j}{\partial \Delta x_i}$ and $\frac{\partial y_j}{\partial \Delta y_i}=0$, or 1 according as j is less than i or otherwise.

5. *Minimum Length enclosing a given Area in Polar Coordinates.*

This problem is evidently equivalent to the previous one of maximizing the area for a given length.

I. For the length to be minimized, and the constant conditions, we have

$$\mathrm{F}=\sum_{i=1}^{n}\sqrt{\Delta r_i^2+r_i^2\Delta\theta_i^2}, \quad . \quad . \quad . \quad . \quad (35)$$

subject to

$$\mathrm{G}\equiv c_1\Big[\sum_{i=1}^{n}\Delta\theta_i-\mathrm{K}\Big]+c_2\sum_{i=1}^{n}\Delta r_i=0\,;$$

$$\mathrm{H}\equiv\sum_{i=1}^{n}\frac{r_i^2\Delta\theta_i}{2}-\mathrm{A}=0. \quad . \quad (36)$$

Here A is evidently the constant area to be enclosed by the minimum F and, as before, the vector condition $\mathrm{G}=0$ specifies that F shall be a definite integral, *i. e.*, between fixed limits and is equivalent to the two scalar conditions $\sum_{=1}^{n}\Delta\theta_i=\mathrm{K}$ and $\sum_{i=1}^{n}\Delta r_i=0$ needed to include various possible *closed* curves.

Here F and H depend exclusively on the $2n$ interval variables Δr_i, $\Delta\theta_i$, since r_i is taken as a function exclusively of the former :

$$r_i=r_0+\sum_{j=1}^{i}\Delta r_j. \quad . \quad . \quad . \quad . \quad . \quad (37)$$

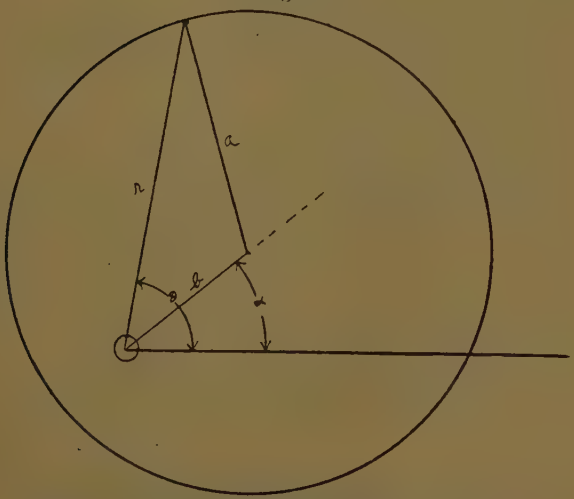
II. Now set the n partial derivatives of $\mathrm{J}=\mathrm{F}-\mathrm{G}+\lambda\mathrm{H}$, with respect to $\Delta\theta_i$ equal to zero, thereby treating the Δr_i's as arbitrary constants. And likewise set equal to zero the partials with respect to Δr_i, treating the $\Delta\theta_i$ as constants:

$$\left.\begin{aligned}\frac{\partial\mathrm{J}}{\partial\Delta\theta_i}&=\frac{r_i^2\Delta\theta_i}{\sqrt{\Delta r_i^2+r_i^2\Delta\theta_i^2}}-c_1+\lambda\frac{r_i^2}{2}=0,\\ \frac{\partial\mathrm{J}}{\partial\Delta r}&=\frac{\Delta r_i}{\sqrt{\Delta r_i^2+r_i^2\Delta\theta_i^2}}+\sum_{j=i}^{n}\frac{r_j\Delta\theta_j^2}{\sqrt{\Delta r_j^2+r_j^2\Delta\theta_j^2}}\\ &\qquad -c_2+\lambda\sum_{j=i}^{n}r_j\Delta\theta_j=0.\end{aligned}\right\}. \quad . \quad (38)$$

Fig. 5.

The interval variables of the length F that encloses a fixed area A.

Fig. 6.



The arc of minimum length that encloses a fixed area A.

The sums in the second equation appear since by (37) $\frac{\partial r_j}{\partial \Delta r_i}=0$ or 1, according as j is less than i or otherwise. Each of these n sets of equations determines n critical constant values of the variables. The former determines n $\Delta\theta_i$'s in terms of n arbitrary Δr_i's and *vice versa* with the latter.

III. Drop the subscripts and interpret (38) as *two* algebraic equations in the *variables* Δr, $\Delta\theta$ subject to (36) and (37) instead of the $2n$ special cases (38).

$$\left.\begin{aligned} \frac{\partial J}{\partial \Delta\theta} &= \frac{r^2\Delta\theta}{\sqrt{\Delta r^2+r^2\Delta\theta^2}} - c_1+\lambda\frac{r^2}{2}=0, \\ \frac{\partial J}{\partial \Delta r} &= \frac{\Delta r}{\sqrt{\Delta r^2+r^2\Delta\theta^2}} + \sum_{i=i}^{n}\frac{r\Delta\theta^2}{\sqrt{\Delta r^2+r^2\Delta\theta^2}} \\ &\qquad -c_2+\lambda\sum_{i=i}^{n} r\Delta\theta=0. \end{aligned}\right\} \quad . \quad (39)$$

Here again, as with (38), the sums start with the particular pair Δr, $\Delta\theta$ (not summed) appearing in the first term. These equations are satisfied by the n pairs Δr_i, $\Delta\theta_i$ that satisfy (38) as well as by infinitely many more.

IV. Now pass to the limit and obtain by elementary calculus

$$\left.\begin{aligned} \frac{\partial J}{\partial d\theta} &= \frac{r^2 d\theta}{\sqrt{dr^2+r^2 d\theta^2}} - c_1+\lambda\frac{r^2}{2}=0, \\ \frac{\partial J}{\partial dr} &= \frac{dr}{\sqrt{dr^2+r^2 d\theta^2}} + \int_{\theta}^{\theta_n}\frac{r\,d\theta}{\sqrt{dr^2+r^2 d\theta^2}}d\theta \\ &\qquad -c_2+\lambda\int_{\theta}^{\theta_n} r\,d\theta=0. \end{aligned}\right\} \quad . \quad (40)$$

Simultaneously with (40), (35) and (36) become the integral limits

$$F=\int\sqrt{dr^2+r^2 d\theta^2}\,;\quad G\equiv\int\left(c_1+c_2\frac{dr}{d\theta}\right)d\theta-c_1K=0\,;$$

$$H\equiv\tfrac{1}{2}\int r^2\,d\theta-A=0. \quad . \quad (41)$$

Eqs. (40) are the new critical conditions, either one of which determines the extremal arc that minimizes the

integral F subject to G=0 and H=0 of (41). Note the *variable* lower limit θ for the integral terms of the above integral differential equation.

Using new constants $c_1=\frac{a^2-b^2}{2a}$ and $\lambda=-\frac{1}{a}$, we can solve the first of (40) for $\frac{d\theta}{dr}$ and rewrite it

$$\frac{d\theta}{dr}=\frac{a^2-b^2+r^2}{r\sqrt{2a^2b^2+2a^2r^2+2b^2r^2-a^4-b^4-r^4}},$$

which integrates to

$$\theta-c=\arcsin\left(\frac{r^2+b^2-a^2}{2br}\right), \quad . \quad . \quad . \quad (42)$$

or

$$\left.\begin{aligned} a^2&=b^2+r^2-2br\sin(\theta-c)\\ a^2&=b^2+r^2-2br\cos(\theta-\alpha)\end{aligned}\right\}\ \alpha=c+\frac{\pi}{2}.$$

Eq. (42) gives the extremal arc that minimizes (41) subject to the constant conditions G=0 and H=0. Evidently it is the polar equation of the circle (see fig. 6) of radius a enclosing area A with origin distant b from the centre along a line making angle α with the initial line. Note that our solution is general or complete in that we have three constants, precisely the right number to uniquely determine the circle. Special cases occur when $a=b$, whereby the origin lies on the circumference : $r=2a\cos(\theta-\alpha)$, and when $b=0$, whereby the origin is at the centre : $r=a$. This solution (42) may be shown to satisfy also the second order equation obtained by differentiating the second eq. (40) with respect to θ so as to remove the integral signs, whereby we obtain Euler's necessary conditions. See (50).

6. *The New Critical Conditions lie at the Foundation of Euler's Equations.*

We may generalize the foregoing problems by applying the same procedure to a more general summation.

Summation Stage I.—

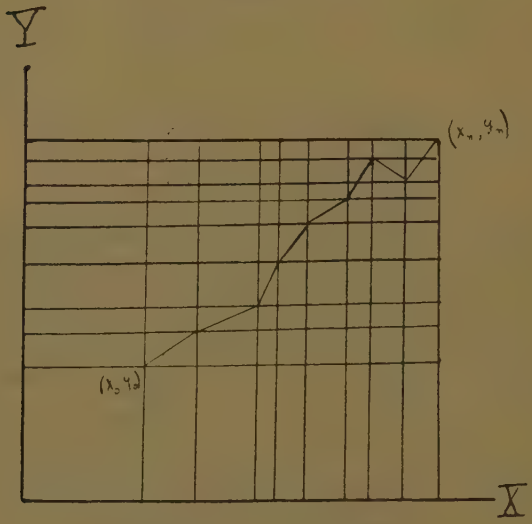
$$\mathrm{F}=\sum_{i=1}^{n} f\left(x_i, y_i, \frac{\Delta y_i}{\Delta x_i}\right)\Delta x_i \quad . \quad . \quad . \quad . \quad (43)$$

is a summation to be maximized or minimized subject to the vector condition

$$G \equiv c_1[\Sigma \Delta x_i - (x_n - x_0)] + c_2[\Sigma \Delta y_i - (y_n - y_0)] = 0, \quad (44)$$

which includes, according to (8), the undetermined multipliers c_1 and c_2 of the Lagrange method, as well as two indicated separate scalar conditions on the Δx_i's and Δy_i's, whereby F becomes a definite sum ranging between the two points (x_0, y_0) and (x_n, y_n). Here the intervals

Fig. 7.



The arbitrary Δx and Δy partitions.

$x_n - x_0$ and $y_n - y_0$ are each partitioned into n successive intervals indicated by fig. 7 in terms of which x_i and y_i are given by

$$x_i = x_0 + \sum_{j=1}^{i} \Delta x_j; \quad \text{and} \quad y_i = y_0 + \sum_{j=1}^{i} \Delta y_j, \quad . \quad (45)$$

which become in the limit as the norms of the Δx and Δy partitions approach zero:

$$x = x_0 + \int_{x_0}^{x} dx, \quad \text{and} \quad y = y_0 + \int_{y_0}^{y} dy. \quad . \quad (45a)$$

Formula (43) is determined by (45) as a function *only* of the $2n$ finite interval variables Δx_i and Δy_i, like (7), (16), (25), (32), and (35) above. Also it is assumed to have like these latter, the usual elementary properties of continuity, etc., so that the operations of Stages I., II., III., and IV. can be performed.

Arithmetic Stage II.—Now form the function $J=F-G$ where the undetermined constants are by (44) included in G, and apply the Lagrange method to obtain the two sets of n critical equations that determine the n critical constants Δx_i in the terms of the n arbitrary Δy_i or *vice versa* (keeping the number n constant) :

$$\left.\begin{aligned} \frac{\partial J}{\partial \Delta x_i} = \frac{\partial F}{\partial \Delta x_i} - \frac{\partial G}{\partial \Delta x_i} = \sum_{j=i}^{n} f_{x_j} \Delta x_j \\ + f_{y_i}{}' \left(- \frac{\Delta y_i}{\Delta x_i{}^2} \right) \Delta x_i + f_i - c_1 = 0, \\ \frac{\partial J}{\partial \Delta y_i} = \frac{\partial F}{\partial \Delta y_i} - \frac{\partial G}{\partial \Delta y_i} = \sum_{j=i}^{n} f_{y_j} \Delta x_j \\ + f_{y_i}{}' - c_2 = 0. \end{aligned}\right\} \quad . \quad . \quad (46)$$

The summation terms appear since by (45) Δx_i is present in every x_j unless j is less than i, that is, $\frac{\partial x_j}{\partial \Delta x_i} = 0$ or 1 according as j is less than i or otherwise. Here f_{x_i}, f_{y_i}, and $f_{y_i}{}'$, are the partial derivatives with respect to the indicated arguments of (43), the last one being taken with respect to $\frac{\Delta y_i}{\Delta x_i} = y_i{}'$. Also we assume that (43) is such a function of Δx_i and Δy_i that (46) may be solved for these $2n$ quantities.

Note that the two sets of n equations (46) refer to *alternative* situations. The first set starts with an *arbitrary* fixed partition of the interval $y_n - y_0$ into $n\,\Delta y_i$'s, and its n equations together with (44) suffice to determine the n critical values of the $n\,\Delta x_i$ plus the undetermined constant c_1. Likewise the second set of (46) together with (44) determines the $n\,\Delta y_i$ (as well as c_2) that are critical for the *arbitrary* partition of $x_n - x_0$ into $n\,\Delta x_i$'s, as indicated by fig. 7. The first set determines the

extreme value of (43) for an arbitrary partition of $y_n - y_0$ and the second for an arbitrary partition of $x_n - x_0$. Obviously these two F's do not in general have the same value, but depend on, and hence are functions of, these arbitrary respective partitions. Thus the set of n pairs of critical constants Δx_i, Δy_i determined by $\frac{\partial J}{\partial \Delta x_i} = 0$ in (46) is not generally the same as that determined by $\frac{\partial J}{\partial \Delta y_i} = 0$, and these two sets of equations are not equivalent.

Algebraic Stage III.—Drop the subscripts from (46) and interpret the result as two equations in the finite *variables* Δx and Δy. Each of these two equations will be satisfied by the two respective sets of n constants in pairs Δx_i and Δy_i determined by (46), as well as by infinitely many more sets corresponding to different arbitrary partitions as per fig. 7 of $y_n - y_0$ (or $x_n - x_0$) into different sets Δy_i (or Δx_i):

$$\left.\begin{aligned} \frac{\partial J}{\partial \Delta x} &= \sum_{j=i}^{n} f_x \Delta x - \frac{\Delta y}{\Delta x} f_{y'} + f - c_1 = 0, \\ \frac{\partial J}{\partial \Delta y} &= \sum_{j=i}^{n} f_y \Delta x + f_{y'} - c_2 = 0. \end{aligned}\right\} \quad . \quad (47)$$

Here Δx and Δy are paired in that they measure intervals between the same (any) two successive points, and the summation extends from this (any) pair to the last pair that reaches to (x_n, y_n). Each of (47) determines the partition indicated by fig. 7 that makes an extremum of (43) where the other partition is kept arbitrary. Thus the first of (47) determines a Δx partition which depends on the arbitrary Δy partition so that the extreme value of (43) will also depend on the Δy partition, as can be indicated by F_m (Δy-partition). Likewise the second of (47) determines an extreme F that depends on the Δx partition: F_m (Δx-partition), where the subscript indicates that F_m is a relative maximum or minimum of this sort.

Calculus Stage IV.—Now pass to the limit as the norm (maximum value of Δx_i or Δy_i) of each arbitrary

partition approaches zero. Then by calculus we have from (46) and (47)

$$\left.\begin{aligned}\frac{\partial J}{\partial dx}&=\frac{\partial F}{\partial dx}-\frac{\partial G}{\partial dx}=\frac{\partial F}{\partial dx}-c_1\\&=\int_x^{x_n} f_x\,dx-y'f_{y'}+f-c_1=0,\\\frac{\partial J}{\partial dy}&=\frac{\partial F}{\partial dy}-\frac{\partial G}{\partial dy}=\frac{\partial F}{\partial dy}-c_2\\&=\int_x^{x_n} f_y\,dx+f_{y'}-c_2=0.\end{aligned}\right\}\quad\ldots\quad(48)$$

Simultaneously, we have from (43) and (44), the integrals

$$F=\int_{x_0}^{x_n} f(x,\,y,\,y')\,dx;\quad G\equiv\int_{x_0}^{x_n}(c_1+c_2y')\,dx-K=0,\quad(49)$$

where K is a brief symbol for the vector $K=c_1(x_n-x_0)+c_2(y_n-y_0)$. Each pair Δx, Δy of (46) or (47) measures the interval between two successive points, and as these approach coincidence the norms of the partitions indicated by fig. 7 approach zero simultaneously, so that the two eqs. (48) each represents the same limit, even though obtained as the limit of the two sets (46) that do not represent in general the same set of critical $\frac{\Delta y_i}{\Delta x_i}$'s nor the same F_m. The first of (48) comes from letting the norm of the arbitrary Δy partition of (47) approach zero, and the second from letting the norm of the Δx partition do it.

Eqs. (48) are the new critical conditions for an extremum of the integral (49) subject to the constant G=0. We may show mathematically that the two eqs. (48) each represents the same limit except where one or the other is non-existent as applied to vertical or horizontal straight lines. Thus take the total derivative of each with respect to x and y:

$$\left.\begin{aligned}\frac{d}{dx}\left(\frac{\partial J}{\partial dx}\right)&=y'\left(\left(f_y-\frac{d}{dx}f_{y'}\right)\right)=0,\\\frac{d}{dy}\left(\frac{\partial J}{\partial dx}\right)&=-\frac{d}{dx}\left(\frac{\partial J}{\partial dy}\right)=f_y-\frac{d}{dx}f_{y'}=0.\\\frac{d}{dy}\left(\frac{\partial J}{\partial dy}\right)&=-\frac{1}{y'}\left(f_y-\frac{d}{dx}f_{y'}\right)=0.\end{aligned}\right\}\quad\ldots\quad(50)$$

Note the factors y' and $1/y'$ which show that (50) includes the Goldschmidt solution for minimum surfaces of revolution. The first of (50) (as of (46), (47), and (48)) includes the horizontal line solution having $\Delta y = dy = 0$ everywhere, whereby the last equation is non-existent. Likewise the last of eqs. (46), (47), (48) and (50) include the vertical line solutions with $\Delta x = dx = 0$, and the first of the equations in this group is non-existent. Eqs. (50) also show how (48) may serve as a foundation for Euler's second-order equations obtained from these by differentiation.

7. *Admissible Arcs and Functions.*

Eqs. (48) and integrals (49) are here assumed to exist and the former are taken to have existing solutions, while the latter include the indicated integrals for any $f(x, y, y')$ as well as any $y(x)$ with its y' and constants of integration, that make the integral exist, whether the arc in question satisfies (48) or not. Thus all arcs and functions $y(x)$ and $f(x, y, y')$ are admissible for the present treatment save those for which (48) and (49) do not exist and serve their purpose as indicated above.

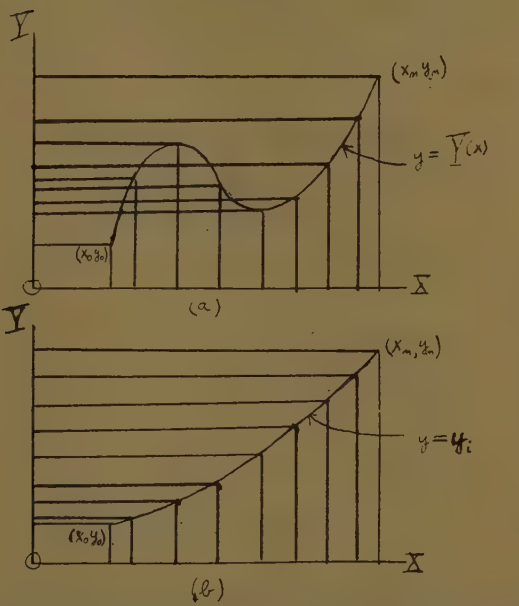
Accordingly these two formulæ (48) and (49) together constitute the mathematical test (the sole test) for admissibility.

All the special cases here treated illustrate arcs and functions admissible because we get results with them. Also the admissible arcs usually associated with Euler's equations as critically necessary for an extremum of (49) are admissible here also, since by (50) the first order (48) include the integral form of Euler's equations, and hence will in general exist and be critical in each case where this is true of Euler's equations. But (48) are more inclusive by virtue of the factors y' and $1/y'$ in (50) and may include multiple valued functions (see fig. 1).

8. *Theorem I.—If there exists a differentiable function* $\phi(x)$ *such that for any positive* ϵ *the inequalities* $|\phi(x_i) - y_i| < \epsilon$ *and* $\left|\phi'(x_i) - \frac{\Delta y_i}{\Delta x_i}\right| < \epsilon$ *can be satisfied by* (a) *choosing a sufficiently fine partition, and simultaneously* (b) *determining the* Δy_i *(and* y_i*) so as to minimize* (43):

$\sum_{i=1}^{n} f\left(x_i, y_i \frac{\Delta y_i}{\Delta x_i}\right) \Delta x_i$, *then* $y=\phi(x)$ *minimizes the integral* (49) : $\int_{x_0}^{x_n} f(x\ y, y')\ dx$.

Fig. 8.



The Δx and Δy partitions for two cases of the summation (43). Note that each has the same Δx partition, whereas the former (a) has its Δy partition determined by the arbitrary $y=Y(x)$, and the latter (b) has it determined by (46) so as to minimize (43).

Hypothesis :

There exist differentiable continuous functions $\phi(x)$ and $f(x, y, y')$ such that :

(a) For every $\epsilon>0$, there exists $a\delta>0$ such that

$$\left| \phi(x_i)-y_i \right| <\epsilon\ ; \quad \text{and} \quad \left| \phi'(x_i)-\frac{\Delta y_i}{\Delta x_i} \right| <\epsilon$$

for every $\Delta x <\delta$; and

(*b*) Every Δy_i (and hence y_i by (45)) is so determined as to minimize (43) :

$$\sum_{i=1}^{n} f\left(x_i, y_i, \frac{\Delta y_i}{\Delta x_i}\right) \Delta x_i \leqq \sum_{i=1}^{n} f\left(x_i, Y_i \frac{\Delta Y_i}{\Delta x_i}\right) \Delta x_i,$$

where the curve $y=Y(x)$ is any differentiable function of x on which the points (x_i, Y_i) lie, and with Y_i as well as x_i and y_i given by (45) in terms of the corresponding increments.

To prove :

$$\int_{x_0}^{x_n} f(x, \phi, \phi')\, dx \leqq \int_{x_0}^{x_n} f(x, Y, Y')\, dx.$$

Proof.—If this inequality is not satisfied, we have

$$\int f(x, \phi, \phi')\, dx - \int f(x, Y, Y')\, dx = t,$$

where t is some positive quantity, and hence the identity :

$$\begin{aligned} t = \int f(x, \phi, \phi')\, dx - \int f(x, Y, Y')\, dx \\ \equiv \left[\int f(x, \phi, \phi')\, dx - \Sigma f(x_i \phi(x_i) \phi'(x_i) \Delta x_i \right] \\ + \left[\Sigma f(x_i \phi(x_i),\ \phi'(x_i)) \Delta x_i - \Sigma f\left(x_i,\ y_i \frac{\Delta y_i}{\Delta x_i}\right) \Delta x_i \right] \\ + \left[\Sigma f\left(x_i,\ y_i \frac{\Delta y_i}{\Delta x_i}\right) \Delta x_i - \Sigma f\left(x_i Y_i \frac{\Delta Y_i}{\Delta x_i}\right) \Delta x_i \right] \\ + \left[\Sigma f\left(x_i Y(x_i) \frac{\Delta Y_i}{\Delta x_i}\right) \Delta x_i - \Sigma f(x_i Y(x_i) Y'(x_i) \Delta x_i \right] \\ + \left[\Sigma f(x_i Y(x_i),\ Y'(x_i)) \Delta x_i - \int f(x, Y, Y')\, dx \right]. \end{aligned} \qquad \ldots \quad (51)$$

The integral of the 1st bracket is by calculus the limit of the sum in this bracket, hence for a sufficiently fine Δx partition we have the inequality :

1*st*. $\left[\int f(x, \phi, \phi')\, dx - \Sigma f(x, \phi(x_i) \phi'(x_i) \Delta x_i \right] < \frac{t}{4}.$

Likewise, with the 5th bracket :

5*th*. $\left[\Sigma f(x_i Y(x_i) Y'(x_i)) \Delta x_i - \int f(x, Y, Y')\, dx \right] < \frac{t}{4}.$

For the 2nd bracket we have, by hypothesis, for every $\epsilon > 0$,

$$\left|\phi(x_i) - y_i\right| < \epsilon \quad \text{and} \quad \left|\phi'(x_i) - \frac{\Delta y_i}{\Delta x_i}\right| < \epsilon,$$

for a sufficiently fine partition. Hence using the continuity of the function $f(x, y, z)$, we have the inequality:

$$2nd. \quad \left[\Sigma f(x_i \phi(x_i) \phi'(x_i))\, \Delta x_i - \Sigma f\left(x_i, y_i \frac{\Delta y_i}{\Delta x_i}\right) \Delta x_i\right] < \frac{t}{4}.$$

For the 4th bracket we have for every $\epsilon > 0$

$$\left|\frac{\Delta Y_i}{\Delta x_i} - Y'(x_i)\right| < \epsilon$$

for a sufficiently fine partition, since $\frac{\Delta Y_i}{\Delta x_i}$ is by (45) the slope of the secant line connecting (x_{i-1}, Y_{i-1}) with (x_i, Y_i) and $Y'(x_i)$ is the slope of the tangent line at x. Hence also, using the continuity of the function $f(x, y, z)$, we have the inequality:

$$4th. \quad \left|\Sigma f\left(x_i, Y_i \frac{\Delta Y_i}{\Delta x_i}\right) \Delta x_i - \Sigma f(x_i, Y_i, Y'(x_i))\, \Delta x_i\right| < \frac{t}{4}$$

for a sufficiently fine partition. Thus the 1st, 2nd, 4th, and 5th brackets of (51) are taken together subject to the inequality $4\epsilon < t$, and since the third bracket is negative or zero by hypothesis, we see that (51) is proved to be incorrect. Hence:

$$\int_{x_0}^{x_n} f(x, \phi, \phi')\, dx - \int_{x_0}^{x_n} f(x, Y, Y')\, dx \leqq 0. \quad \text{Q.E.D.}$$

Theorem II.—If the conditions of Theorem I. are satisfied, and if for every partition the Δy_i and y_i which minimize (43) *satisfy the conditions* (46), *then $\phi(x)$ satisfies conditions* (48).

Hypothesis:

Given $f(x, y, y')$ with continuous f_y and $f_{y'}$ and $y = \phi(x)$, continuous differentiable functions which satisfy (*a*) and

(b) of Theorem I., and also (c) such that for every partition, the Δy_i (and hence y_i by (45)) which minimize (43): $\sum_{i=1}^{n} f\left(x_i, y_i, \frac{\Delta y_i}{\Delta x_i}\right) \Delta x_i$, satisfy the conditions

$$\frac{\partial J}{\partial \Delta y_i} = \sum_{j=i}^{n} f_{y_j} \Delta x_j + f_{y'_i} - c_2 = 0. \quad . \quad . \quad . \quad (46)$$

To prove that $y=\phi(x)$ satisfies

$$\frac{\partial J}{\partial dy} = \int_x^{x_n} f_y\, dx + f_{y'} - c_2 = 0. \quad . \quad . \quad . \quad (48)$$

Proof.—By hypothesis (a) we have for every $\epsilon>0$

$$\left|\phi(x_i)-y\right| < \epsilon \quad \text{and} \quad \left|\phi'(x_i) - \frac{\Delta y_i}{\Delta x}\right| < \epsilon,$$

for a sufficiently fine partition. Hence for f_y and $f_{y'}$ continuous functions of these variables we have:

For every $\epsilon>0$, no matter how small,

$$\left|\int_x^{x_n} f_y\, dx - \sum_{j=i}^{n} f_{y_j} \Delta x_j\right| + \left|f_{y'} - f_{y'_i}\right| < \epsilon,$$

for a sufficiently fine partition. If we replace the indexed terms according to (46), we have

$$\left|\int_x^{x_n} f_y\, dx + f_{y'} - c_2\right| < \epsilon$$

for a sufficiently fine partition. Since ϵ may be taken arbitrarily small, the arguments $\phi(x)$ and $\phi'(x)$ of the left number are thus shown to be the ordinate and slope of a solution of (48), since this is the equation we get by setting this last left member equal to zero.—Q. E. D.

Likewise, these two theorems may be proved for the first of eqs. (46) and (48) by exchanging places of the Δx and Δy partitions in the line of reasoning, the latter being now arbitrary and the *former* now so chosen as to minimize (43) and satisfy the *first* of eqs. (46).

Likewise, also, we can show that (48) are the critical equations for a maximum of (49) in which proof, the left member and middle bracket of (51) would have their signs reversed.

An existence theorem to show that the limits $y=\phi(x)$ with slope $\phi'(x)$ exist and also proof of the existence of the

critical equations (48) and the integral (49) is not included in the present article, which deals primarily with method. It will readily be seen that (48) and (49) and certain solutions $y=\phi(x)$ of the former do exist for all the special problems here treated. Also an examination of all the problems in Bolza's ' Variationsrechnung ' indicates that (48) and (49) exist for all of these and that (48) has appropriate solutions for an extremum of (49), although exceptional non-existent cases may be noted. Thus for Bolza's prob. 28, p. 148, the integral (49) is $\mathrm{F}=\int_0^{\pi} y'^2\cos^2 x\,dx$, which has $y=\tan x$ for its extremal arc (taking $b=1$ and $c=0$ in (57) below). For $x=\pi/2$, however, this arc is non-existent and hence inadmissible, and likewise with the integral (49) for the limits in question. For the limits 0 to $\pi/4$, however, this integral exists and becomes unity, a minimum compared with F for other existing arcs such as $y=4x/\pi$ for which (49) becomes

$$\mathrm{F}=16/\pi^2\int\cos^2 x\,dx=(2\pi+4)/\pi^2>1.$$

Hence this particular F and arc $y\ (x)$ are admissible except when $x=\pi/2$ is included, in which case they are non-existent.

9. *The Fundamental or Delta Process for the Calculus of Variations.*

The four stage delta process illustrated above in sections 1 to 6 may be summarized.

Summation Stage I.—Set up an approximation to a desired integral limit as a summation or function of n pairs (in two dimensions) of variables Δx_i and Δy_i only :

$$f=\sum_{i=1}^{n} f\left(x_i,\ y_i\ \frac{\Delta y_i}{\Delta x_i}\right)\Delta x_i, \quad . \quad . \quad . \quad . \quad (43)$$

where $\quad x_i=x_0+\sum_{j=1}^{i}\Delta x_j, \quad$ and $\quad y_i=y_0+\sum_{j=1}^{i}\Delta y_j.$

Likewise employ the constant condition to specify the direction and magnitude of the space range, being the vector equivalent of the two indicated scalar conditions

$$\mathrm{G}=c_1\left[\sum_{i=1}^{n}\Delta x_i-(x_n-x_0)\right]+c_2\left[\sum_{i=1}^{n}\Delta y_i-(y_n-y_0\right]=0.$$

$$. \quad . \quad . \quad (44)$$

Note that its two partial derivatives with respect to Δx_i and Δy_i are always c_1 and c_2.

Arithmetic * *Stage II.*—Form the function $J=F-G$ and set equal to zero its indicated $2n$ partial derivatives, thereby determining by the Lagrange method n pairs of critical constants Δx_i and Δy_i that make an extremum of F subject to $G=0$, as solutions of either set of n equations :

$$\left.\begin{aligned} \frac{\partial J}{\partial \Delta x_i} &= \frac{\partial F}{\partial \Delta x_i} - \frac{\partial G}{\partial \Delta x_i} = \frac{\partial F}{\partial \Delta x_i} - c_1 = 0, \\ \frac{\partial J}{\partial \Delta y_i} &= \frac{\partial F}{\partial \Delta y_i} - \frac{\partial G}{\partial \Delta y_i} = \frac{\partial F}{\partial \Delta y_i} - c_2 = 0. \end{aligned}\right\} \quad (46)$$

Algebraic Stage III.—Drop the subscripts from Δx_i and Δy_i and thereby obtain two equations in these *variables* satisfied by the n pairs determined by each of the two previous sets of n equations (46) as well as other such sets for any other positive integral n :

$$\left.\begin{aligned} \frac{\partial J}{\partial \Delta x} &= \frac{\partial F}{\partial \Delta x} - \frac{\partial G}{\partial \Delta x} = \frac{\partial F}{\partial \Delta x} - c_1 = 0, \\ \frac{\partial J}{\partial \Delta y} &= \frac{\partial F}{\partial \Delta y} - \frac{\partial G}{\partial \Delta y} = \frac{\partial F}{\partial \Delta y} - c_2 = 0. \end{aligned}\right\} \quad . \quad . \quad (47)$$

The significance of these two equations in the variables Δx and Δy becomes obvious if we consider that dropping the subscripts from the previous set of Stage II. amounts to nothing more than departing from the two particular sets of n pairs Δx_i and Δy_i there treated, as n ranges through all positive integral values. Thus J, F, and G are the same functions of Δx and Δy in (47) as they are of Δx_i and Δy_i in (43) and (46), and hence admit differentiation by hypothesis in the same way as indicated. The change from (46) to (47) may be viewed as merely one of notation such as always is involved when an equation in variables is used for brevity to represent a set of equations that refer to a set of constants. It involves, incidentally, the generalization that enters by the inclusion of other sets of constants besides the original

* Called the arithmetic stage because once the partial derivatives are set equal to zero we deal with n pairs of unique critical constants rather than with two equations in variables as in (47), and hence the *transition* from Stage II. to Stage III. is like the transition from arithmetic to algebra.

ones—a valid generalization, because the new sets of constants for different values of n are just as admissible as the original set and cannot be even distinguished from them so long as we use a general n.

Calculus Stage IV.—Pass to the limit as the norms of the Δx and Δy partitions approach zero, and thereby obtain the critical differential equations (48) that determine the extremum of F, which now becomes the definite integral (49) subject to the integral condition G=0, which specifies the limits of the definite integral

$$\left.\begin{aligned} \frac{\partial J}{\partial dx} &= \frac{\partial F}{\partial dx} - \frac{\partial G}{\partial dx} = \frac{\partial F}{\partial dx} - c_1 = 0, \\ \frac{\partial J}{\partial dy} &= \frac{\partial F}{\partial dy} - \frac{\partial G}{\partial dy} = \frac{\partial F}{\partial dy} - c_2 = 0. \end{aligned}\right\} \quad . \quad (48)$$

Eqs. (48) represent what their form indicates, the symbols being given precise meaning by Working Rules I. and II. of section 10 below. They indicate differentiation with respect to differentials, operating on J=F−G, as a function of dx and dy only, just as the summations (43) and (44) are functions of Δx_i and Δy_i only, and (45 a) gives a familiar formula for the dependence of x and y on their differentials. Hence we see that in treating J=F−G as a function of dx and dy only, we do not depart radically from the familiar calculus view-point which considers an integral as the limit of a sum. Even the form is in accord with custom, for it is usual to interpret the differentials of the independent variables dx and dy as equal to the corresponding interval variables Δx and Δy, giving special attention to what occurs as they approach zero. Recall, too, the proof by Theorem II. that (48), the limiting form of (46) and (47), is likewise critical for an extremum of (49) the limit of (43). This is shown to follow from the way the arguments y_i and $\frac{\Delta y_i}{\Delta x_i}$ approach $\phi(x_i)$ and $\phi'(x_i)$ as the Δx partition approaches zero.

Thus the change from (47) to (48) may for the case in hand be viewed as merely one of notation,as was the case with the change from (46) to (47). With this Calculus Stage IV., given by (48), however, it is a matter of *emphasis* to give special attention to what occurs as $\Delta x = dx$ and $\Delta y = dy$ approach zero.

In case of isoperimetric problems, we have the additional condition like F to be set up in Stage I., $\mathrm{H}\left(x_i, y_i, \frac{\Delta y_i}{\Delta x_i}\right)=0$. The remaining Stages II. to IV. are precisely as above, except that F is to be replaced by $\mathrm{F}+\lambda\mathrm{H}$, where λ is an undetermined constant. The Lagrange multiplier method holds for any number of conditions, so that for this additional one we have $\mathrm{J}=\mathrm{F}+\lambda\mathrm{H}-\mathrm{G}$, as is illustrated in sections 4 and 5 above. Then instead of (48) we have for Stage IV.

$$\left.\begin{aligned} \frac{\partial \mathrm{J}}{\partial dx} &= \frac{\partial \mathrm{F}}{\partial dx}+\lambda\frac{\partial \mathrm{H}}{\partial dx}-c_1=0, \\ \frac{\partial \mathrm{J}}{\partial dy} &= \frac{\partial \mathrm{F}}{\partial dy}+\lambda\frac{\partial \mathrm{H}}{\partial dy}-c_2=0. \end{aligned}\right\} \quad . \quad . \quad . \quad (52)$$

These are the new critical equations for an extremum of F, a definite integral with limits determined by $\mathrm{G}=0$, and subject to the isoperimetric * condition $\mathrm{H}=0$.

10. *Working Rules applicable to* single *integration for differentiation of a definite integral with respect to differentials.*

The above delta process gives uniform results expressible briefly in terms of two workng rules whose proof rests on the above delta process itself whose results they express.

Working Rule (*I*).—The partial derivative of an integral with respect to a differential that appears *explicitly* under the integral sign is given by (*a*) dropping the integral sign and (*b*) differentiating with respect to this differential, treating it as an ordinary finite variable.

Working Rule (*II*.).—The partial derivative with respect to a differential (say dx), of an integral that depends *implicitly* on this differential by virtue of explicit dependence on the corresponding integral variable (x), as by ($45\,a$) above, is given by (*a*) differentiating under the integral sign with respect to this integral variable (x) and (*b*) setting as *lower* limit, that same or other integral *variable* (x or y) with respect to which the integration is to be carried out.

These rules are seen to apply to any case of single integration. Thus in all cases above we have set equal

* Illustrated by (29), (34), and (40) above.

to zero the partial derivatives of a definite integral J with respect to its differentials, regardless of its special form as a combination of F and G and sometimes H. Thus the rules for differentiation of a definite integral with respect to its differentials are hereby established, although their significance is not touched upon here except where J has the above special forms and is applied to problems in the calculus of variations.

Note that these two working rules do *not* apply in the above form to double integration, since (29) above is not obtainable by their direct application. Very likely, however, the procedure may be extended and rules established for differentiation of double and triple integrals with respect to their differentials.

11. *Euler's rule for the isoperimetric case is obtained by differentiating the new critical equations* (52).

If we let $h(x, y, y')$ be the integrand for the isoperimetric constant integral $H = \int h(x, y, y')dx - k = 0$ of (52), we may carry out the indicated operations, using working rules (I.) and (II.) and obtain

$$\left.\begin{aligned} \frac{\partial J}{\partial dx} &= \int_x^{x_n} (f_x + \lambda h_x)\,dx + (f + \lambda h) - y'(f_{y'} + \lambda h_{y'}) - c_1 = 0, \\ \frac{\partial J}{\partial dy} &= \int_x^{x_n} (f_y + \lambda h_y)\,dx + f_{y'} + \lambda h_{y'} - c_2 = 0, \end{aligned}\right\} \quad . \quad . \quad . \quad (53)$$

which are the new critical conditions for the extremum of F subject to the isoperimetric condition $H = 0$. To obtain Euler's rule, we have only to take the total derivative of each of (53) with respect to x and y:

$$\left.\begin{aligned} \frac{d}{dx}\left(\frac{\partial J}{\partial dx}\right) &= y'\left[f_y + \lambda h_y - \frac{d}{dx}(f_{y'} + \lambda h_{y'})\right] = 0, \\ \frac{d}{dy}\left(\frac{\partial J}{\partial dx}\right) &= -\frac{d}{dx}\left(\frac{\partial J}{\partial dy}\right) = f_y + \lambda h_y - \frac{d}{dx}(f_{y'} + \lambda h_{y'}) = 0, \\ \frac{d}{dy}\frac{\partial J}{\partial dy} &= -\frac{1}{y'}\left[f_y + \lambda h_y - \frac{d}{dx}(f_{y'} + \lambda h_{y'})\right] = 0. \end{aligned}\right\} \quad . \quad . \quad . \quad (54)$$

Eqs. (53) are illustrated by (34) and (40) above in cartesian and polar coordinates applied to Dido's problem. Here,

as with (50), the first and third of eqs. (54) include horizontal and vertical lines as admissible solutions because of the factors y' and $1/y'$.

12. *Theorem III.—The minimizing arc for the integral* $F=\int(y'^2/u')\,dx$ *has the form* $y=bu+c$, *where* u *is any continuous function of* x *only, having first derivatives everywhere in the region under consideration.*

Given u any continuous function of x only, having first derivatives everywhere between the limits of integration of the definite integral *

$$F=\int_{x_0}^{x_n} f(x,\ y,\ y')\,dx=\int_{x_0}^{x_n}(y'^2/u')\,dx=\lim_{\Delta x\to 0}\sum_{i=1}^{n}\frac{\Delta y_i^2}{\Delta u_i};\quad .\ .\ .\quad (55)$$

where the primes indicate differentiation with respect to x. Along with this, we have the vector restriction

$$G=c_1\left[\int_{x_0}^{x_n}dx-(x_n-x_0)\right]+c_2\left[\int_{x_0}^{x_n}dy-(y_n-y_0)\right]=0.$$

We minimize this integral by applying the second of eqs. (48) according to Working Rule (I.) of section 10, the second being chosen because for it the integral term vanishes, since by (55) $f_y=0$. The critical equation is

$$\frac{\partial J}{\partial dy}=f_{y'}-c_2=2y'/u'-c_2=0,\quad .\ .\ .\quad (56)$$

which integrates to

$$y=bu+c,\quad .\ .\ .\ .\ .\ .\ .\quad (57)$$

where $b=c_2/2$, and hence (55) becomes

$$F=b\int_{y_0}^{y_n}dy=b(y_n-y_0).$$

The advantage of this theorem is that it supplies a definite integral that is minimized by any extremal arc

* This problem is an extension of the conditional minimum of $F=\frac{x^2}{a}+\frac{y^2}{b}$ subject to $x+y=G$, which yields the critical relation $\frac{x}{a}=\frac{y}{b}$. Note that F so determined is a minimum, *i. e.*, a case of least squares.

whatsoever of the type here considered and given by (57) with u and u' continuous within the limits of integration. There is, of course, no limit to the number of choices of u and u' continuous at points where u'' (the second derivative) is non-existent. At every such point, however, Euler's second-order equations are non-existent and hence not necessary conditions for the minimum of (55), which may still exist and be minimized by an existing (57). Thus an indefinite number of places are indicated where Euler's second-order equations are unnecessary, whereas (48) and (56) remain existent and critical.

As illustrations take $y=bu+c$, with u given by

$$u=ax+x^3\sin(1/x) \text{ at } x\neq 0, \text{ with } u=0 \text{ at } x=0, \quad (57\,a)$$

and

$$u=ax+(\pi/2-x)^3\sin(\tan x) \text{ at } x\neq\pi/2, \text{ with } u=\pi a/2 \text{ at } x=\pi/2. \quad (57\,b)$$

The second derivative u'' is non-existent for these expressions at $x=0$ and $x=\pi/2$ respectively, since the former involves the limit $\cos(1/x)$ as x approaches zero *, and the latter the limit $\cos(\tan x)$ as x approaches $\pi/2$. The first derivative, however, exists at these points and is everywhere different from zero if we choose "a" large enough. Hence (55) and (56) exist and are critical at these points where Euler's second-order equations are non-existent and unnecessary. When we use (57 a) and (57 b) and the indicated limits of integration, we have for these two existing cases of (55)

$$\mathrm{F}=b\int_{-2/\pi}^{+2/\pi} y'\,dx=\frac{4ab^2}{\pi}, \quad \text{and} \quad \mathrm{F}=b\int_0^{\pi} y'\,dx=\pi ab^2.$$

13. *Applications of the New Working Rules.*

(1) Brachistochrone: The curve of quickest descent from (x_0, y_0) to (x_n, y_n) is derived † by minimizing

$$\mathrm{F}=\int_{x_0}^{x_n}\sqrt{\frac{dx^2+dy^2}{y-\alpha}},$$

subject to

$$\mathrm{G}\equiv\int_{x_0}^{x_n}(c_1+c_2y')\,dx-\mathrm{K}=0.$$

* Townsend, E. J. 'Functions of Real Variables,' pp. 153, 155.
† Bliss, G. A., 'Calculus of Variations,' p. 45.

Rule (I.) section (10) suffices to give the critical condition if we use the first equation of (48) because

$$f=\sqrt{\frac{1+y'^2}{y-\alpha}},$$

$$J=F-G,$$

and hence $f_x=0$ and (48) is a first-order equation

$$\frac{\partial J}{\partial dx}=\frac{dx}{\sqrt{(y\alpha)-(dx^2+dy^2)}}-c_1=0,$$

or

$$\frac{1}{\sqrt{(y-\alpha)(1+y'^2)}}=c_1.$$

This is the well-known condition that integrates to the necessary cycloid.

(2) Catenary. For the curve that gives a minimum surface of revolution we have to minimize

$$F=\int_{x_0}^{x_n} y\sqrt{dx^2+dy^2},$$

subject to

$$G=\int_{x_0}^{x_n}(c_1+c_2y')\,dx-K=0.$$

Here again $f_x=0$ since the integrand is $f=y\sqrt{1+y'^2}$, so that we employ the first of (48) and differentiate partially with respect to dx, as per Working Rule (I.),

$$J=F-G,$$

$$\frac{\partial J}{\partial dx}=\frac{y\,dx}{\sqrt{dx^2+dy^2}}-c_1=0, \quad \text{or} \quad \frac{y}{\sqrt{1+y'^2}}=+c_1=b.$$

This is the well-known equation whose solution is the catenary, being the basic critical condition (48) applied to the case in hand.

(3) Geodesics on a Sphere. We have to minimize the integral

$$F=a\int\sqrt{d\theta^2+\sin^2\theta\,d\phi^2},$$

subject to

$$G=\int_{\theta_0}^{\theta_n}\left(c_1+c_2\frac{d\phi}{d\theta}\right)d\theta-K=0.$$

Here θ is evidently the polar angle and ϕ the meridian angle of polar coordinates measured on this sphere of radius a. As before we pick out the differential $(d\phi)$ which has no corresponding variable (ϕ) as the one to vary during the partial differentiation process. Then we employ Working Rule (I.) and according to (48) differentiate with respect to $d\phi$ instead of dx; using $J=F-G$,

$$\frac{\partial J}{\partial\, d\phi}=\frac{a\sin^2\theta\, d\phi}{\sqrt{d\theta^2+\sin^2\theta\, d\phi^2}}-c_2=0.$$

This is the desired critical condition that integrates to $\sqrt{a^2-c_2^2}\sin(\phi-\alpha)=c_2\cot\theta$, or in cartesian coordinates

$$\sqrt{a^2-c_2^2}(z\cos\alpha-y\sin\alpha)=c_2x,$$

which is a plane through the origin which intersects the sphere in great circles. $\dfrac{\partial J}{\partial\, d\theta}$ gives a cumbersome integral term with variable lower limit like (24) and (40) above because the integral variable θ requires the application of Working Rule (II.) in the process of differentiation.

(4) Geodesics on a Cone of Revolution. Taking the origin at the vertex of the cone and α as the constant polar angle for points on the cone, we have in polar coordinates

$$F=\int\sqrt{dr^2+r^2\sin^2\alpha\, d\phi^2},$$

and

$$G=\int_{r_0}^{r_n}\left(c_1+c_2\frac{d\phi}{dr}\right)dr-K=0.$$

As before, we differentiate with respect to the differential whose corresponding integral variable ϕ is absent from J and hence avoid the integral term in our basic condition (48):

$$J=F-G,$$

$$\frac{\partial J}{\partial\, d\phi}=\frac{r^2\sin^2\alpha\, d\phi}{\sqrt{dr^2+r^2\sin^2\alpha\, d\phi^2}}-c_2=0,$$

which integrates quite readily to the surface

$$r\cos\{(\phi-b)\sin\alpha\}=c,$$

whose intersection with the cone $\theta=\alpha$ gives the geodesic in question.

(5) Newton's Problem. The surface of revolution of minimum resistance involves minimizing.

$$\mathrm{F}=\int_{x_0}^{x_n}\frac{yy'^3}{1+y'^2}\,dx=\int_{x_0}^{x_n}\frac{y\,dy^3}{dx^2+dy^2},$$

subject to

$$\mathrm{G}=\int_{x_0}^{x_n}(c_1+c_2y')\,dx-\mathrm{K}=0.$$

Use (48) and differentiate with respect to dx since the integrand is free from x, and obtain, as per Working Rule (I.),

$$\mathrm{J}=\mathrm{F}-\mathrm{G},$$

$$\frac{\partial \mathrm{J}}{\partial\,dx}=\frac{-2y\,dy^3\,dx}{(dx^2+dy^2)^2}-c_1=0,$$

which may be written

$$y=\frac{a}{2}\left(\frac{(1+p^2)^2}{p^3}\right),\quad \text{where } p=\frac{dy}{dx}.$$

Differentiate with respect to x and set the left member equal to p, thereby getting a first-order equation in x and p, which integrates to

$$x=\frac{a}{2}\left(\log p+\frac{1}{p^2}+\frac{3}{4p^4}\right)+c.$$

These last two equations give the solution in well-known * parametric form, showing two arbitrary constants $c_1=-a$ and c.

(6) Hamilton's Principle. The laws of mechanics are given by the critical equations (48) for an extremum of an integral between point events, whose integrand, $\mathrm{L}=\mathrm{T}-\mathrm{V}$, is the difference between kinetic and potential energy.

$$\mathrm{F}=\int_{t_0}^{t_n}(\mathrm{T}-\mathrm{V})\,dt,$$

$$\text{where}\quad \mathrm{T}=\frac{m}{2}\left[\left(\frac{dx}{dt}\right)^2+\left(\frac{dy}{dt}\right)^2+\left(\frac{dz}{dz}\right)^2\right],$$

subject to the vector

$$\mathrm{G}=\int_{t_0}^{t_n}(-c_1+c_2x'+c_3y'+c_4z')\,dt-\mathrm{K}=0.$$

* Bolza, 'Vorlesungen über Variationsrechnung,' §54, Equations (69), (70).

Note that the dimensions of c_1 are L/T times the dimensions of the c's of G for the space intervals, and that the opposite sign of c_1 is required by (58) and (60) if the energy and momentum and the c's are all to be positive. This is to be expected, since this G is a constant interval between point events and plays a role similar to the invariant interval $dl^2=c^2\,dt^2-dx^2-dy^2-dz^2$ of the Theory of Relativity. Take V independent of time and apply Working Rule (I.) to the integral J=F−G :

$$\frac{\partial J}{\partial\, dt}=T-V-dt\left[\, m\left(\frac{dx^2}{dt^3}+\frac{dy^2}{dt^3}+\frac{dz^2}{dt^3}\right)\right]$$

$$+c_1=T-V-2T+c_1=0, \quad \text{or} \quad c_1=E=T+V, \quad . \quad (58)$$

which is the conservation of energy condition with $c_1=E$, the energy constant. The total derivative of (58) with respect to t gives

$$\frac{d}{dt}\left(\frac{\partial J}{\partial\, dt}\right)=-\frac{dx}{dt}\left[\frac{\partial V}{\partial x}+m\frac{d^2x}{dt^2}\right]$$

$$-\frac{dy}{dt}\left[\frac{\partial V}{\partial y}+m\frac{d^2y}{dt^2}\right]-\frac{dz}{dt}\left[\frac{\partial V}{\partial z}+m\frac{d^2t}{dt^2}\right]=0. \quad (59)$$

Eqs. (58) and (59) are the new conditions (48) and Euler's equations (50) applied to the Hamiltonian integral, the former representing conservation of energy and the latter the force equations of mechanics, as is well known.

Now differentiate J=F−G partially with respect to dx instead of dt, using Rule (II.) as well as (I.), since V depends on the integral variable x,

$$\frac{\partial F}{\partial\, dx}=c_2+\frac{\partial J}{\partial\, dx}=m\frac{dx}{dt}-\int_t^{t_n}\frac{\partial V}{\partial x}\,dt=c_2=M_x. \quad (60)$$

This is the conservation of momentum condition. Its time derivative is

$$\frac{d}{dt}\left(\frac{\partial J}{\partial\, dx}\right)=\frac{\partial V}{\partial x}+m\frac{d^2x}{dt^2}=0. \quad . \quad . \quad . \quad (61)$$

Eqs. (60) and (61) are again (48) and (50) applied to the case in hand, and represent conservation of momentum and the force equations of mechanics according to Euler's equations. Likewise with the y and z components.

If V, and hence L=T−V, depends explicitly on the

time, the energy equation (58) would have an integral term like that of (60),

$$-\int_t^{t_n} \frac{\partial V}{\partial t} dt$$

according to the Working Rule (II.). The total derivative of this term with respect to t being $\frac{\partial V}{\partial t}$, we have this additional term in (59)

$$\frac{d}{dt}\left(\frac{\partial J}{\partial \dot{dt}}\right) = \frac{\partial V}{\partial t} - \left\{\frac{dx}{dt}\left(\frac{\partial V}{dx} + m\frac{d^2x}{dt^2}\right)\right.$$

$$\left. + \frac{dy}{dt}\left(\frac{\partial V}{\partial y} + m\frac{d^2y}{dt^2}\right) + \frac{dz}{dt}\left(\frac{\partial V}{\partial z} + m\frac{d^2z}{dt^2}\right)\right\} = 0. \quad (62)$$

This is a " continuity " equation for energy rather than mass, and its dimensions are " power." It says that the time rate of change of the potential energy equals the resultant time rate of doing work by all the forces.

(7) Principle of Least Action. The new method brings out clearly the distinction between Hamilton's Principle and the Principle of Least Action. The former involves a variation between fixed point events, the latter involves a variation between fixed space interval coordinates where the time of transit may vary. This latter comes from assuming the energy constant $E=T+V$ given by (58). Thus we drop the time component of the vector condition

$$G_t = \int_{t_0}^{t_n} dt - (t_n - t_0) = 0$$

and have the vector sum of only the other three conditions for the three *space* intervals

$$G = \int_{x_0}^{x_n} \left(c_2 + c_3\frac{dy}{dx} + c_4\frac{dz}{dx}\right) dx - K = 0,$$

where $K = c_2(x - x_0) + c_3(y - y_0) + c_4(z_n - z_0)$.

Take the action integral

$$F = \int_{t_0}^{t_n} 2T\, dt,$$

where T is the kinetic energy as above, and assume (58) and introduce E=T+V, and obtain

$$F=2\int_{t_0}^{t_n}\sqrt{T(E-V)}\,dt$$

as the action integral in Jacobi form that is required to remain stationary, *i. e.*, satisfy the new critical conditions (48) subject to G above. Thus

$$\left.\begin{aligned} &J=F-G,\\ &\frac{\partial J}{\partial dx}=m\frac{dx}{dt}\sqrt{\frac{E-V}{T}}-\int_t^{t_n}\frac{\partial V}{\partial x}\sqrt{\frac{T}{E-V}}\,dt-c_2=0,\\ &\text{or}\qquad m\frac{dx}{dt}-\int_t^{t_n}\frac{\partial V}{\partial x}\,dt=+c_2=M_x, \end{aligned}\right\} \quad . \ . \ . \quad (63)$$

where the simplification comes from setting the radicals equal to unity according to (58). Equation (63) is the momentum equation, identical with (60), that leads to the force equation (61), which is the Euler condition for the extremum of F herein. Likewise for the y and z components. We note the basic difference between Hamilton's Principle and that of Least Action. The former treats of a variation between fixed point events, and the new critical conditions are the conservation of energy (58) or the continuity of energy (62), as well as conservation of momentum (60). On the other hand, the principle of Least Action assumes (58), a *result* of Hamilton's Principle as a starting point, and the continuity (62) of a variable energy is outside its scope. It treats of a variation between fixed space coordinates only, and the new basic condition (63) is the conservation of momentum equation. Both principles, of course, give the force equations of mechanics (Euler's equations) as the total derivative with respect to t of the new critical conditions that represent conservation of energy and momentum, as is done for the general case where (48) gives rise by differentiation to Euler's equations (50).

The writer is specially indebted to Professor G. Y. Rainich of Ann Arbor for very valuable criticism and discussions, and also special thanks are due to Professor J. E. Powell of East Lansing.

XII. *The Effect of Air Coupling in Acoustic Insulation by means of Elastic Supports.* *By* A. H. Davis, *D.Sc.*, *and* A. E. Knowler, *M.Sc.*, *A.Inst.P.*, *Physics Department, National Physical Laboratory, Teddington, Middlesex* *.

IN isolating floors or foundations from vibration due to attached machinery it is customary to support the machinery upon rubber pads or upon a concrete block so supported. Again, floors may be isolated to some extent from vibration due to a person walking thereon by covering the structural floor with a finishing floor of, say, concrete "floating" on rubber pads. In either case the principle is to arrange the pads to be so small and elastic that the natural frequency of oscillation of the load upon the elastic supports is as low as possible, preferably in the infrasonic range below 30 cycles per second, and certainly low compared with the frequencies of the vibrations against which it is desired to insulate.

The object of this note is to point out that if the interspace between the concrete finish and the structural floor is enclosed at the edges and is only moderately thick the forces transmitted to the structural floor by the air film exceed those transmitted *viâ* the pads.

For, let a rigid slab of mass m lie flat upon a set of elastic supports of such stiffness that a vertical displacement of the slab from its equilibrium position calls into play a restoring force sx. The natural frequency of vertical oscillation of the slab on the supports is then $n/2\pi$, where $n^2=s/m$. If an oscillatory force $F \cos \omega t$ is applied to the upper side of the slab the fraction ϵ of the maximum force transmitted to the floor is given by

$$\epsilon = \sqrt{\frac{\frac{4\Delta^2}{n^2}+\frac{n^2}{\omega^2}}{\frac{4\Delta^2}{n^2}+\frac{n^2}{\omega^2}+\frac{\omega^2}{n^2}-2}}\ \dagger,$$

* Communicated by the Authors.

† C. R. Sodeburg, Elec. J. xxi. p. 161 (1924), gives this result in a different form.

where the internal damping of the system is represented by the decay factor Δ.

If now the interspace of depth l be enclosed at the edges the enclosed air will exert a further elastic restraint upon the motion of the upper slab. The instantaneous excess pressure p set up in the air by a displacement x of the slab will be $c^2\rho x/l$, where ρ is the density of the enclosed air and c the velocity of sound in it. The total force exerted by the air on the floor underneath the slab will thus be $A\rho c^2 x/l$ * if A is the surface area of the slab. Calling this $s_1 x$, where $s_1 = \rho c^2 A/l$, the total stiffness of the elastic supports and enclosed air is $S = s + s_1$. The natural frequency and transmissibility of the system under the new conditions with enclosed air in the interspace are obtained by writing S for s in the equations above.

To take a special case, let the slab be of 5 cm. concrete, 10,000 sq. cm. in area, and carry such a load that the total weight is 100 kilograms. This slab, supported at each corner on a 2·5 cm. cube of soft rubber, would have, in the absence of air coupling, a natural frequency (N) equal to about 8 cycles per second. From this we may infer that s ($=4\pi^2 N^2 m$) would be $0{\cdot}25 \times 10^9$ c.g.s. units. With the interspace air enclosed s_1 would have a value of approximately $15 \times 10^9/l$ c.g.s. units, where l is expressed in cm. Thus, with a gap of 60 cm. $s_1 = s$—that is, the stiffness (s_1) due to the air equals that (s) due to the rubber; however, with a gap of 1 cm. $s_1 = 60s$—that is, the stiffness due to the air would be about sixty times as great as that due to the rubber.

Assuming the forcing vibration to be of frequency 160 cycles per second, the transmissibility in the case of a 1 cm. layer of air would be 0·18, as against 0·0025 in the case of rubber supports with no confined air in the interspace. This amount corresponds to a 15 decibel reduction in the transmitted force, instead of the 52 db. theoretically attainable with rubber coupling alone. Figures for other thicknesses of interspace air are included in the following table. The deleterious effect of insufficient separation is obvious.

* If an enclosure, in changing from volume V to $V - dV$, compresses the contained air adiabatically, the pressure P_0 changes to $P_0 + p$, where $p/P_0 = \gamma dV/V$, where γ is the ratio of specific heats of air. This yields $p = \rho c^2 (dV/V)$.

Any considerable damping of the system by viscous forces would involve amendments of the calculated insulation, lower gains being obtained in each case. In some experiments described below, however, the damping was estimated from the form of the resonance curve when the floating floor was driven at various frequencies by an unbalanced motor and was not sufficient ($\Delta/n=0{\cdot}04$) to affect the calculations given in the table by more than 1 db.

The general tendency revealed by the above figures can be illustrated experimentally by supporting on a normal concrete floor the rubber-insulated slab of concrete

Insulation afforded by " floating " floor of 5 cm. concrete, 10,000 sq. cm. in area, carrying a load such that the total weight is 100 kilograms, and supported near each corner on a 2·5 cm. cube of soft rubber. Interspace air confined.

Separation between structural and floating floors.	Calculated insulation against 160~ alternating force.	Observed insulation against rubber-shod " footstep " machine.
cm.	db.	phons.
1	15	16
2	21	21
5	29	29

referred to above, and placing upon it a mechanically driven rubber-shod hammer delivering blows corresponding to fairly heavy footfalls. The interspace air should be confined. By suitable threaded sockets one can arrange continuous adjustment of the extent to which the rubber isolators are let into the concrete block, and can thus allow continuous variation of the interspace depth. A person listening below the floor then observes considerable increase in the loudness of the transmitted noise as the thickness of the interspace is reduced. It was, in fact, through an experiment of this kind that the importance of the air coupling was first noticed. Some measurements of the insulation afforded against impacts by a floating slab of the size mentioned above are given in the last column of the table. The slab was a central section of a large floating floor, and was insulated from the

remainder by felt. The hammer machine delivered four blows per second. It is true that the exciting force was not harmonic, but most of the sound was of the frequency of the fundamental mode of vibration of the slab—a frequency found by vibrograph observations to be about 160 cycles per second, the value taken for the frequency of the exciting force in the above calculations. The equivalent loudness of the sound transmitted to the room below was measured by means of a subjective noise meter of the Barkhausen type.

It may be mentioned that with the 2 cm. separation increasing the number of rubber supports from 4 to 32, or using rubber pads of greater stiffness, had no appreciable effect on the measured insulation. This is consistent with the calculation that the coupling due to the interspace air is greater even than that due to the 32 pads of rubber.

The close agreement in the table between the results (in decibels) calculated for insulation against harmonic alternating force and those (in phons) observed for insulation against "footsteps" is partly accidental. Some differences between the experimental conditions and the conditions assumed in the calculations have been mentioned above; but in addition to these the calculations relate to complete confinement of the interspace air, whereas in practice the confinement was only partial—an effect which would *lower* the insulation in decibels below the calculated value given in the table. On the other hand, a change of, say, 15 decibels in the intensity of a 160-cycle tone causes a change of possibly 20 phons in its equivalent loudness—a circumstance which tends to *raise* the value of the loudness change in phons above the numerical value of the calculated intensity change in decibels given in the table.

The accuracy to which the various effects compensate each other is not known, but the table expresses the experimental fact that the spacing is important, and indicates that the order of the importance of spacing is predictable from calculations of the stiffness of the interspace air.

It is clear that when a floating floor is used for insulating the main structural floor against vibration set up by footsteps or similar disturbances coupling due to the interspace air can be of considerable importance and can limit the advantage obtainable.

XIII. *On Second-Order Adjoint Difference Systems.*

By Voris V. Latshaw *.

In considering by the methods of Lagrange the motion of a vibrating string of beads one deals with the following recurrence relation

$$A_{r-1}+CA_r+A_{r+1}=0.$$

This relation is slightly generalized in the self-adjoint difference equation

$$L(u)\equiv u(x+1)+\lambda p(x)u(x)+u(x-1)=0, \quad . \quad . \quad (1)$$

where x takes on the values $(a, a+1, \ldots, b-1)$ with both $p(x)$ and λ real. Discussions of such second-order difference equations together with various types of boundary conditions are to be found in the literature *. The present note is devoted chiefly to a discussion of second-order adjoint difference systems. The definition of such systems is to be formulated by methods analogous to those usually associated with differential systems. For a Lagrange relation we use the finite sum

$$\sum_{x=a}^{b-1}[v(x)L(u)-u(x)L(v)]=\pi(u, v), \quad . \quad . \quad . \quad . \quad (2)$$

where

$$\pi(u, v)=u(b)v(b-1)-v(b)u(b-1)-u(a)v(a-1)+v(a)u(a-1), \quad . \quad (3)$$

a skew-symmetric form. $\pi(u, v)$ can be transformed into canonical form. To this end we select arbitrarily four linearly independent forms,

$$U_i=a_{i1}u(a-1)+a_{i2}u(a)+b_{i1}u(b-1)+b_{i2}u(b), \ (i=1, 2, 3, 4),$$

and write

$$\pi(u, v)=U_1V_4+U_2V_3+U_3V_2+U_4V_1.$$

The two systems

$$L(u)=0 \ldots . \ (4\,a) \qquad U_1=U_2=0 \ldots . \ (4\,b)$$

$$L(v)=0 \ldots . \ (5\,a) \qquad V_1=V_2=0 \ldots . \ (5\,b)$$

thus developed we shall call adjoint difference systems, and, if equivalent, self-adjoint difference systems.

* Communicated by the Author.

† T. Fort, 'Quarterly Journal of Pure and Applied Mathematics,' No. 181, p. 1 (1914); H. Levy and E. A. Baggott, Phil. Mag. (7) xviii. p. 177 (1934).

2. Let the pair of functions (y_1, y_2) be a fundamental set of solutions of $(4\,a)$. The set (w_1, w_2) related to this set by the equations

$$w_1(x)y_1(x)+w_2(x)y_2(x)=0, w_1(x)y_1(x+1)+w_2(x)y_2(x+1)=1$$

is said to be adjoint to it. That w_1 and w_2 reduce $L(w)$ to zero is readily verified. It follows, then, that we can write

$$\left.\begin{aligned} w_1(x)&=c_{11}y_1(x)+c_{12}y_2(x).\\ w_2(x)&=c_{21}y_1(x)+c_{22}y_2(x).\end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (6)$$

To each of these equations we apply the operator θ, $(\theta f(x)=f(x+1)$. We are then in a position to show by direct calculation that the matrix C of the substitution (6) is skew-symmetric. This statement holds also for systems of higher order.

3. If λ is so chosen that system (4) is incompatible, system (5) also has only a trivial solution. Let $f(x)$ be a solution of (4) and (y_1, y_2) a fundamental set for $L(y)=0$. By substituting in (2) we have

$$U_1(y_1)V_4(f)+U_2(y_1)V_3(f)=0,$$

$$U_1(y_2)V_4(f)+U_2(y_2)V_3(f)=0.$$

Hence $V_1(f)=V_2(f)=V_3(f)=V_4(f)=0$, which implies $f\equiv 0$. Under these circumstances there exists a uniquely determined function $G(x, t)$ defined for $(a-1\leqslant x\leqslant b)$, $(a\leqslant t<b)$ with the properties $L[G(i, j)]=\delta_{ij}$ and $U_1(G)=U_2(G)=0$. Let $H(x, t)$ be such a function for (5). In the sum (2) by setting $u=G(x, t_1)$ and $v=H(x, t_2)$, one finds $G(t_2, t_1)=H(t_1, t_2)$, $(a\leqslant t_1, t_2<b)$. If the system (4) is self-adjoint, G will be symmetric. Conversely, symmetry of G implies self-adjointness. To prove the converse we let $f(x)$ defined for $(a-1\leqslant x\leqslant b)$ denote any function satisfying $(5\,b)$ and form the sum

$$\sum_{x=a}^{b-1}[fL(G)-GL(f)]=\pi(G, f)=0.$$

We find

$$f(t)=\sum_{x=a}^{b-1}G(x, t)L[f(x)].$$

Let an arbitrary function $F(x)$ be defined as follows:

$$F(x)=\sum_{=a}^{b-1}G(x, t)L[f(t)].$$

As a consequence one has

$$\mathrm{L}[f(x)]=\mathrm{L}[\mathrm{F}(x)], \qquad (a\leqslant x<b),$$
$$\mathrm{F}(x)=f(x), \qquad (a-1\leqslant x\leqslant b).$$

Since $\mathrm{F}(x)$ satisfies the equations (4 b), $f(x)$ also does. That any function satisfying (4 b) also satisfies (5 b) can be demonstrated similarly. Thus we have shown that symmetry of $\mathrm{G}(x, t)$ is necessary and sufficient for self-adjointness of system (4) under the assumption of incompatibility.

4. A criterion better suited to our purpose and equivalent to the one just written is available. Let $u(x)$ and $v(x)$ be any pair of functions defined for $(a-1\leqslant x\leqslant b)$ satisfying (4 b). We can write

$$u(t)=\sum_{x=a}^{b-1}\mathrm{G}(t, x)\mathrm{L}[u(x)], \quad v(t)=\sum_{x=a}^{b-1}\mathrm{G}(t, x)\mathrm{L}[v(x)].$$

The substitution of these functions in (2) gives

$$\sum_{t,\ x=a}^{b-1}\{\mathrm{G}(t, x)-\mathrm{G}(x, t)\}\mathrm{L}[u(t)]\mathrm{L}[v(x)]=\pi(u, v).$$

If G is symmetric, $\pi(u, v)$ vanishes identically. Conversely, in view of the arbitrary nature of u and v, $\pi(u, v)\equiv 0$ implies symmetry of G. This criterion for self-adjointness of system (4) remains valid without the assumption of incompatibility. We shall show that if $\pi(u, v)\equiv 0$ for every pair of functions (u, v) satisfying (4 b), then a non-trivial solution $f(x)$ of (4) is also a solution of system (5). To do this we set $v=f$ and $u=u_1$, u_2 in (2). These substitutions provide relations which insure that $\mathrm{V}_1(f)=\mathrm{V}_2(f)=0$. Moreover, it can be shown that there do not exist two linearly independent solutions of (5) unless the same can be said for system (4).

We now make use of the criterion just developed in determining the explicit form of the conditions for self-adjointness of system (4). By rearranging and renaming the terms in (4 b) one can so write them

$$q_{11}u_1+q_{12}u_2+\dot{q}_{11}\dot{u}_1+\dot{q}_{12}\dot{u}_2=0,$$
$$q_{21}u_1+q_{22}u_2+\dot{q}_{21}u_1+\dot{q}_{22}\dot{u}_2=0,$$

that the determinant $\mathrm{A}_1=|\,q_{ij}\,|$ is non-vanishing. Let the skew-symmetric form (3) be rewritten in terms of the newly defined quantities

$$\pi(u, v)=\mathrm{C}(u_1v_2-u_2v_1)+\mathrm{C}'(u_1\dot{v}_2-\dot{u}_2\dot{v}_1)+\sum_{r,\ s=1}^{2}\mathrm{C}''_{rs}(u_r\dot{v}-\dot{u}_sv_r).$$

By making use of the boundary conditions we can write this expression as follows :

$$\pi(u,v)=\mathrm{T}(\dot{u}_1\dot{v}_2-\dot{u}_2\dot{v}_1)/\mathrm{A}_1.$$

For (4) to be self-adjoint T must vanish, that is,

$$\mathrm{T}=\mathrm{CA}_2+\sum_{r=1}^{n}\{\mathrm{C}''_{r1}\,\mathrm{D}_r(2)-\mathrm{C}''_{r2}\mathrm{D}_r(1)\}+\mathrm{A}_1\mathrm{C}'=0,$$

in which $\mathrm{D}_r(\mathrm{i})$ denotes the determinant A_1 with q_{kr} replaced by $\dot{q}_{ki}$ and $\mathrm{A}_2=|\,\dot{q}_{ij}\,|$. To illustrate the results let us assume the non-singular matrix to be

$$\mathrm{A}_1=\begin{vmatrix} a_{11} & b_{11} \\ a_{21} & b_{21} \end{vmatrix}.$$

We have

$$\mathrm{T}=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}-\begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix}=0. \quad . \quad . \quad (7)$$

Other possible assumptions with regard to A_1 lead to the same result. Boundary conditions satisfying this restriction can be written down at once. For example :

$$a_{11}u(a-1)+u(a)+b_{12}u(b)=0,$$
$$b_{12}u(a-1)+u(b-1)+b_{22}u(b)=0.$$

For another type we select a set of Sturmian boundary conditions

$$a_{11}u(a-1)+a_{12}u(a)=0,$$
$$b_{11}u(b-1)+b_{12}u(b)=0.$$

Clearly these conditions satisfy (7). On the other hand, if both boundary conditions pertain to point "*a*" or point "*b*" (4) cannot be self-adjoint.

XIV. *The Sound Insulation of Single and Complex Partitions.* *By* J. E. R. Constable, *M.A., Ph.D., B.Sc., and* G. H. Aston, *M.A., Ph.D., B.Sc., Physics Department, National Physical Laboratory, Teddington, Middlesex* *.

Abstract.

Results of measurements made at the National Physical Laboratory of the sound insulating value of a number of

* Communicated by Dr. G. W. C. Kaye, O.B.E., M.A.

partitions are given together with curves which exhibit the correlation between the sound insulation and weight of single homogeneous partitions. The results are discussed and certain principles which should be adopted in sound insulating constructions are deduced.

IN the course of the last few years the insulation afforded against air-borne sound by a number of partitions of various types has been measured at the National Physical Laboratory. Some of this work has been carried out at the request of various companies, the remainder being for experimental purposes. In the following pages an account is given of the results obtained, and of some of the conclusions which can be drawn from them.

The measurements were made upon partitions of dimensions 5 feet 2 inches by 3 feet 10 inches. The partitions were clamped against a reveal built into an aperture connecting two otherwise acoustically isolated rooms lined with sound absorbent material. In one of the rooms approximately pure tones from a loud-speaker were directed obliquely towards the aperture, the disposition of the apparatus being such that there was no appreciable acoustical reaction upon the loud-speaker when the aperture was closed by the test partition. In the other room measurements of the transmitted sound were made with a traversing microphone both with and without the test partition in position in the aperture. The sound insulating value of the partition, expressed in decibels, is ten times the common logarithm of the ratio which the sound energy transmitted through the open aperture bears to that transmitted through the test partition.

A more detailed description of the test rooms and the method has been given by A. H. Davis and T. S. Littler in an earlier paper *. The parabolic reflector and loud-speaker unit which was then used to obtain a parallel beam of sound is now replaced by a large moving coil diaphragm type loud-speaker.

From the point of view of sound insulation it is possible to divide partitions into two groups, one of which may be termed single partitions and the other complex partitions.

* A. H. Davis and T. S. Littler, Phil. Mag. iii. p. 177 (1927).

The distinction is somewhat arbitrary, but, for the purpose of this paper, the following are included in the first group :—Single sheets, solid panels constructed from one material, and single walls of bricks or building blocks (plastered or unplastered). This group may be further subdivided into single homogeneous partitions such as building board, glass, and solid brick walls, and single inhomogeneous partitions such as walls of hollow tiles. Partitions other than single partitions are termed complex.

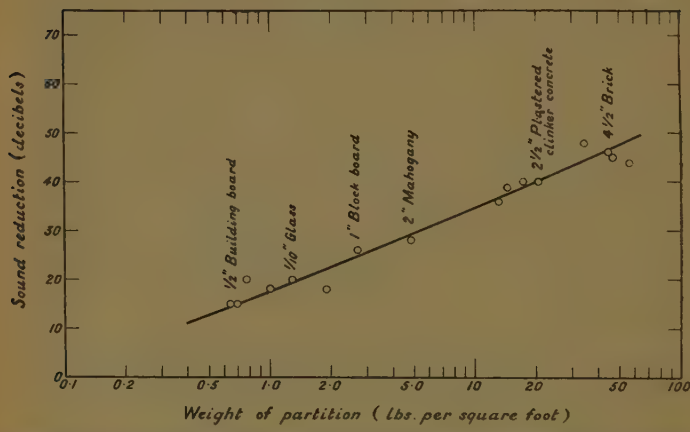
The basis of this classification is that when sound falls upon a single partition the vibrations of the surface upon which it is incident are transmitted directly to the other surface. In a complex partition, on the other hand, the vibrations have to pass through some intermediate material, *e. g.*, air, cork, or felt.

Measurements made at the National Physical Laboratory and other laboratories * have shown that there is a very good correlation between the sound insulation of single homogeneous partitions of widely varying material and their weight per square foot. This correlation is exhibited in the curves shown in figs. 1 to 4 which were obtained from measurements of such partitions in the small sound transmission rooms at the National Physical Laboratory. Figs. 1 to 3 were obtained by averaging results for the following groups of test frequencies : 200 and 300 ; 500, 700, and 1000 ; 1600 and 2000 cycles per second. Fig. 4 was obtained by averaging results for all the seven frequencies.

These curves can be regarded as representing the sound insulation which the average single homogeneous partition of a given weight may be expected to have when tested under the conditions of these measurements. They are for this reason very useful in assessing the effectiveness of any partition construction. For example, a complex type of partition is of special value for sound insulation only in so far as it has an insulation appreciably greater

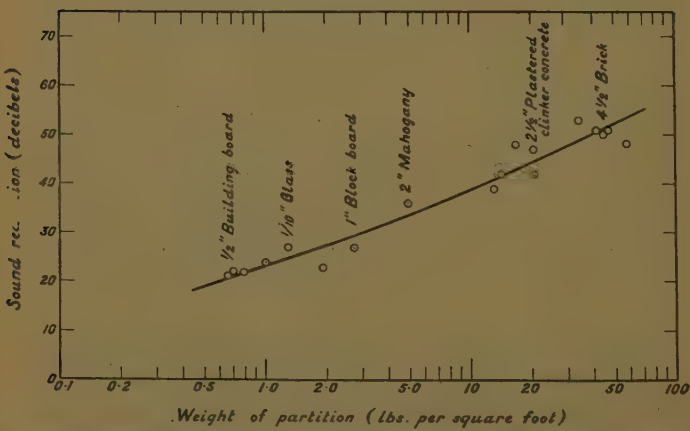
* A. H. Davis, Phil. Mag. xv. p. 309 (1933). ; Chrisler and Snyder, Bureau of Standards J. of Research, ii. p. 541 (1929) ; E. Meyer, *Preuss. Akad. Ber. Phys. Math. Klasse*, p. 171 (1931) ; Hofbauer and Benz, *Zeits. der Osterr. Ingenieur u. Arch. Vereines*, Heft 9/10 (1934) · V. Knudsen, 'Architectural Acoustics,' p. 300 ; H. Kreuger and J. H. Sager, *Ingeniörs Vetenskaps Akad. Handlingar*, no. 132 (1934) ; P. E. Sabine, 'Acoustics and Architecture,' p. 262 ; R. Berger, *Forschg. Ing-Wes.* iii. p. 200 (1932) ; E. Weisbach, 'Dissertation.' (1910).

Fig. 1.



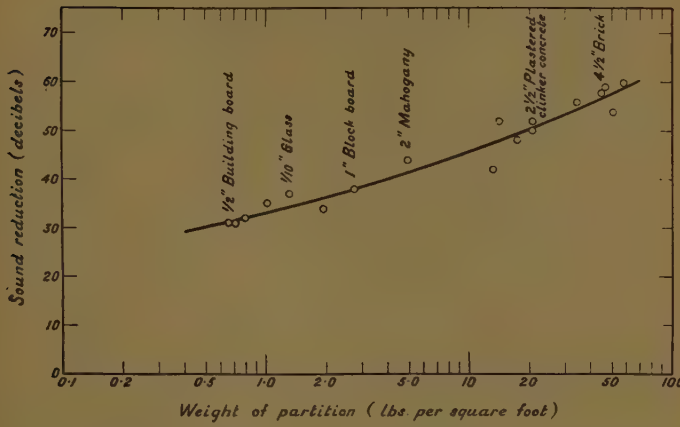
Relation between sound reduction in decibels and weight per square foot of single homogeneous partitions. Sound reduction averaged for frequencies 200 and 300 cycles per second.

Fig. 2.



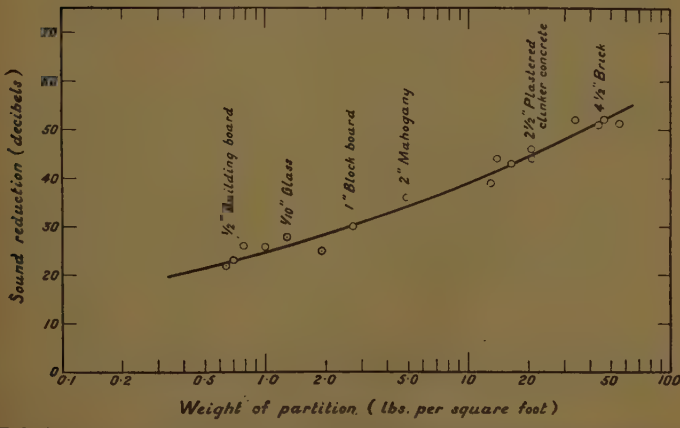
Relation between sound reduction in decibels and weight per square foot of single homogeneous partitions. Sound reduction averaged for frequencies 500, 700, and 1000 cycles per second.

Fig. 3.



Relation between sound reduction in decibels and weight per square foot of single homogeneous partitions. Sound reduction averaged for frequencies 1600 and 2000 cycles per second.

Fig. 4.



Relation between sound reduction in decibels and weight per square foot of single homogeneous partitions. Sound reduction averaged for frequencies 200, 300, 500, 700, 1000, 1600, and 2000 cycles per second.

than that of the average single homogeneous partition of the same weight.

As shown by the scatter of the points through which the curves are drawn, single homogeneous partitions may be expected to show divergences of several decibels from the curves. The divergences are probably mainly due to resonances of the partitions tested and are largest when results at a single frequency are considered ; experience shows that in this case the mean divergence is about 3 db., the maximum being 10 db. In this paper results averaged for bands of frequencies are used since the divergences are thereby reduced, *e. g.*, in the case of the curves shown in figs. 1, 2, 3, the mean divergences lie between 1 and 2 db., the maximum being 6 db.

In Tables I. and II. are given the sound insulation of a number of single and complex partitions selected from among those measured at the National Physical Laboratory during the last few years. Practically all the partitions were enclosed in frames of $1\frac{1}{2}$ to 2-inch wood, exceptions being building boards, sheets of glass, and complex partitions nos. 36, 45–55, 60, 71, and 80. In the first column of each table is a brief description of the construction of the partition ; the second and third columns give the weight in pounds per square foot (an average for the partition and the frame) and the thickness respectively. The other columns give the sound insulation averaged for the frequencies specified with, in brackets, the sound insulation from figs. 1–4 for an average single homogeneous partition of the same weight. The last column contains a sound insulation figure obtained by averaging the results for all the test frequencies. This is included, as it provides a useful method of roughly comparing partitions although, since the insulation of some complex partitions varies more rapidly with frequency than that of others, it is not always reliable when such partitions are being compared. The rate of variation of insulation with frequency is, however, about the same for all single homogeneous partitions, and hence in their case the average for the entire frequency range has more significance.

Discussion.

I. *Single Homogeneous Partitions.*

The correlation between weight and sound insulation has led to various attempts being made to calculate

Table I. A.

Single Homogeneous Partitions.

Partition number.	Description of test partition.			Average sound reduction for frequencies (cycles per second).			
	Construction.	Weight (lbs./sq. ft.).	Thickness (inches).	200 and 300.	500, 700, and 1000.	1600 and 2000.	200, 300, 500, 700, 1000, 1600, and 2000.
				db.	db.	db.	db.
1.	Essex board	0·61	$\frac{3}{16}$	14*	21†	29‡	—
2.	Ankarboard insulating board ...	0·64	$\frac{1}{2}$	15 (14)	21 (21)	31 (31)	22 (22)
3.	Treetex insulating board	0·69	$\frac{1}{2}$	15 (15)	23 (22)	31 (31)	23 (23)
4.	Donnacona insulating board ...	0·70	$\frac{1}{2}$	11*	22†	34‡	—
5.	Insulwood panel..	0·75	$\frac{1}{2}$	16*	21†	32§	—
6.	Lloyd wall board .	0·77	$\frac{1}{2}$	20 (16)	25 (22)	32 (32)	26 (23)
7.	Plywood	1·0	$\frac{3}{8}$	18 (18)	24 (24)	35 (34)	26 (25)
8.	21-oz. window glass	1·3	$\frac{3}{32}$	20 (19)	27 (25)	37 (35)	28 (26)
9.	Balsa wood panel..	1·9	$1\frac{7}{8}$	18 (22)	23 (27)	34 (36)	25 (28)
10.	Gaboon-faced block board....	2·7	1	26 (25)	27 (30)	38 (38)	30 (31)
11.	Plate glass	3·8	$\frac{1}{4}$	28 (28)	33 (32)	45 (40)	35 (32)
12.	Mahogany	4·9	2	28 (30)	36 (34)	44 (42)	36 (34)
13.	Wood-fibre slabs plastered on both faces	13·1	2	36 (37)	39 (41)	42 (48)	39 (41)
14.	Cellular anhydrite blocks	14·3	3	39 (38)	42 (42)	52 (48)	44 (42)
15.	Cellular anhydrite blocks plastered on both faces with Pioneer plaster	16·8	$3\frac{1}{4}$	40 (39)	48 (43)	50 (50)	46 (43)
16.	3-inch Eonit blocks plastered on both faces ..	20·4	$3\frac{7}{8}$	40 (40)	47 (45)	50 (51)	46 (45)
17.	$2\frac{1}{2}$-inch clinker concrete, plastered on both faces	20·7	$3\frac{1}{2}$	40 (40)	47 (45)	52 (51)	44 (45)

* For 300 cycles per second only.
† Mean for 500 and 1000 cycles per second.
‡ For 1600 cycles per second only.
§ For 2000 cycles per second only.

TABLE I. A (*cont.*).

Partition number.	Description of test partition.			Average sound reduction for frequencies (cycles per second).			
	Construction.	Weight (lbs./sq. ft.).	Thickness (inches).	200 and 300.	500, 700, and 1000.	1600 and 2000.	200, 300, 500, 700, 1000, 1600, and 2000.
				db.	db.	db.	db.
18.	Lignocrete composition bricks, lime plastered on both faces...	34·0	5⅜	48 (44)	53 (49)	56 (54)	52 (50)
19.	4½-inch Fletton bricks, lime mortar	41·0	4½	38*	51 (49)	53‡	—
20.	4½-inch Fletton brick, plastered on both faces (lime mortar, lime and sand plaster)	44·0	5½	46 (47)	50 (51)	58 (56)	51 (52)
21.	4½-inch Fletton brick, plastered on both faces (cement mortar, hard plaster) ..	46·0	5½	45 (47)	51 (51)	59 (57)	52 (52)
22.	Reinforced concrete	50·8	4	45*	50†	58§	—
23.	8-inch panel of Portland cement and "foamed" slag, rendered on one face with cement and sand and plastered on the other with Sirapite	56·8	9⅜	44 (49)	48 (54)	60 (58)	51 (54)

Table I. B.

Single Inhomogeneous Partitions.

Partition number.	Description of test partition.			Average sound reduction for frequencies (cycles per second).			
	Construction.	Weight (lbs./sq. ft.).	Thickness (inches).	200 and 300.	500, 700, and 1000.	1600 and 2000.	200, 300, 500, 700, 1000, 1600, and 2000.
				db.	db.	db.	db.
24.	Thatchboard plastered on both faces	12·4	—	29*	35†	49§	—
25.	Plycrete partition blocks plastered on both faces ..	16·4	3½	38 (39)	45 (43)	48 (50)	43 (43)
26.	Hollow tiles, lime plaster on both faces	21·0	4¼	42 (41)	42 (45)	48 (51)	44 (45)
27.	3-inch hollow clay blocks, glazed on one face and plastered on the other	22·4	3⅜	42 (41)	48 (45)	48 (51)	46 (45)
28.	Fosalsil blocks ...	22·5	—	37*	46†	54§	—
29.	3-inch hollow clay blocks, plastered on both faces ..	23·2	3⅞	40 (42)	46 (46)	48 (52)	45 (46)
30.	4-inch hollow clay blocks, plastered on both faces ..	26·5	4¾	41 (42)	46 (46)	48 (52)	45 (47)
31.	Terra-cotta partition blocks	27·6	—	39*	46†	58§	—
32.	Hollow tiles, lime plaster on both faces	27·7	4¼	44 (43)	50 (47)	47 (53)	47 (48)
33.	Bimol hollow blocks plastered on both faces ..	31·8	—	45*	49†	54‡	—
34.	Fletton cellular bricks	31·8	—	40*	48†	56§	—
35.	4-inch Cranham blocks plastered on both faces ..	37·0	6¼	44 (46)	52 (48)	56 (54)	51 (49)

Table II.
Complex partitions.

Partition number.	Description of test partition.			Average sound reduction for frequencies (cycles per second).			
	Construction.	Weight (lbs./sq. ft.).	Thickness (inches).	200 and 300.	500, 700, and 1000.	1600 and 2000.	200, 300, 500, 700, 1000, 1600, and 2000.
				db.	db.	db.	db.
36.	Two sheets of ½-inch fibre board separated by 2-foot air-space ..	1·5	25	33 (20)	46 (26)	65 (35)	48 (27)
37.	Two sheets of aluminium separated by air-space of about 2 inches	1·6	2	10 (21)	26 (26)	52 (36)	29 (27)
38.	Idem ; interspace filled with cork feathers	1·7	2	10 (22)	26 (26)	47 (36)	27 (28)
39.	Idem ; interspace filled with Kapok	1·8	2	10 (22)	30 (27)	50 (36)	30 (28)
40.	Idem ; interspace filled with Kapok blanket ...	1·8	2	10 (22)	29 (27)	52 (36)	30 (28)
41.	Idem ; interspace filled with eel-grass.....	1·8	2	13 (22)	27 (27)	48 (36)	29 (28)
42.	A layer of ½-inch fibre board on each side of a frame of 2-inch wood. 1-inch air-space between layers ...	1·9	2	17 (22)	35 (27)	56 (36)	36 (28)
43.	6 layers of Alfol special foil (0·007 mm. gauge) attached to wood frame in space between ¼-inch 3-ply wood and ⅛-inch compressed wood board	2·1	2	21*	28†	52§	—
44.	Thermolux 3-ply glass	2·2	$\frac{1}{16}$	22 (24)	28 (28)	40 (37)	30 (29)
45–55	Double window of 21-oz. glass at following spacings :—						

TABLE II. (*cont.*).

Partition number.	Description of test partition.			Average sound reduction for frequencies (cycles per second).			
	Construction.	Weight (lbs./sq. ft.).	Thickness (inches).	200 and 300.	500, 700, and 1000.	1600 and 2000.	200, 300, 500, 700, 1000, 1600, and 2000.
				db.	db.	db.	db.
	8 inches	2·6	$8\frac{3}{8}$	40 (24)	52 (29)	66 (38)	53 (30)
	7 ,, ...	—	$7\frac{3}{16}$	36	47	62	48
	6 ,, ...	—	$6\frac{3}{16}$	36	50	67	51
	5 ,, ...	—	$5\frac{3}{16}$	35	49	64	49
	3 ,, ...	—	$3\frac{3}{16}$	28	52	64	49
	2 ,, ...	—	$2\frac{3}{16}$	23	49	64	46
	1 ,, ...	—	$1\frac{3}{16}$	20	42	65	42
	$\frac{3}{4}$,, ...	—	$\frac{13}{16}$	18	41	64	41
	$\frac{1}{2}$,, ...	—	$\frac{11}{16}$	22	36	56	38
	$\frac{1}{4}$,, ...	—	$\frac{7}{16}$	23	32	52	35
	$\frac{1}{8}$,, ...	—	$\frac{5}{16}$	24	27	50	33
56.	A layer of $\frac{1}{2}$-inch fibre board on each side of a frame of 2-inch wood. 3-inch air-space between layers ...	2·6	4	23 (24)	43 (29)	62 (38)	42 (30)
57.	Two sheets of aluminium with interspace filled with Zonolite ..	3·2	2	22 (26)	30 (30)	46 (39)	32 (32)
58.	A layer of $\frac{1}{2}$-inch Insulwood on one side of 4-inch by 2-inch studding at about 14-inch centres	3·5	$4\frac{1}{2}$	16 (27)	25 (31)	35 (40)	25 (32)
59.	A layer of $\frac{3}{4}$-inch Insulwood on one side of 4-inch by 2-inch studding at about 14-inch centres	3·9	$4\frac{3}{4}$	14 (28)	26 (32)	36 (40)	25 (32)
60.	A layer of $\frac{7}{16}$-inch B.P. insulating board on each side of 4-inch by 2-inch staggered studding with layer of $\frac{1}{2}$-inch Balsam wool between the studs	4·2	6	26 (28)	51 (32)	75 (41)	51 (33)

TABLE II. (*cont.*).

Partition number.	Description of test partition.			Average sound reduction for frequencies (cycles per second).			
	Construction.	Weight (lbs./sq. ft.).	Thickness (inches).	200 and 300.	500, 700, and 1000.	1600 and 2000.	200, 300, 500, 700, 1000, 1600, and 2000.
				db.	db.	db.	db.
61.	A layer of ½-inch Insulwood on each side of 4-inch by 2-inch studding at about 14-inch centres	4·3	5	27 (28)	40 (32)	58 (41)	41 (33)
62.	A layer of ½-inch fibre board on each side of 4-inch by 2-inch studding at 12-inch centres....	4·5	5	24 (29)	42 (33)	56 (41)	41 (34)
63.	A layer of ½-inch L.W. brand fibre board on each side of 4-inch by 2-inch studding at 14-inch centres	4·7	5	31 (29)	46 (34)	59 (42)	46 (34)
64.	A layer of ½-inch fibre board on each side of 4-inch by 2-inch studding at 15-inch centres....	4·9	5	36 (30)	48 (34)	56 (42)	47 (34)
65.	A layer of ¾-inch Insulwood on each side of 4-inch by 2-inch studding at about 14-inch centres	5·2	5½	30 (30)	37 (34)	50 (42)	39 (35)
66.	Two layers of ½-inch Insulwood on one side and one layer of ½-inch Insulwood on the other side of 4-inch by 2-inch studding at about 14-inch centres	5·2	5½	30 (30)	42 (34)	62 (42)	44 (35)
67.	1-inch matchboarding on 4-						

TABLE II. (*cont.*).

Partition number.	Description of test partition.			Average sound reduction for frequencies (cycles per second).			
	Construction.	Weight (lbs./sq. ft.).	Thickness (inches).	200 and 300.	500, 700, and 1000.	1600 and 2000.	200, 300, 500, 700, 1000, 1600, and 2000.
				db.	db.	db.	db.
	inch by 2-inch studding at 12-inch centres....	5·4	5	28 (30)	24 (34)	31 (42)	27 (35)
68.	A layer of ½-inch L.W. brand fibre board on each side of 4-inch by 2-inch staggered studding, each set of studs at 14-inch centres.	5·5	6	27 (30)	42 (34)	61 (43)	43 (35)
69.	Two layers of ¾-inch Insulwood on one side and one layer of ¾-inch Insulwood on the other side of 4-inch by 2-inch studding at about 14-inch centres	6·2	6¼	26 (31)	45 (35)	61 (44)	44 (36)
70.	Staggered 4-inch by 2-inch studding having a single layer of ½-inch Insulwood on one side and on the other two layers of ½-inch Insulwood separated by 2-inch by ¾-inch fillets. Each set of studs on 14-inch centres	7·3	7	38 (33)	47 (36)	68 (44)	50 (37)
71.	Two sheets of ¼-inch plate glass separated by a 5¾-inch air-space	7·5	6¼	44*	57†	73§	—
72.	Staggered 4-inch by 2-inch studding having a single layer of ¾-inch Insulwood						

TABLE II. (*cont.*).

Partition number.	Description of test partition.			Average sound reduction for frequencies (cycles per second).			
	Construction.	Weight (lbs./sq. ft.).	Thickness (inches).	200 and 300.	500, 700, and 1000.	1600 and. 2000.	200, 300, 500, 700, 1000, 1600, and 2000.
				db.	db.	db.	db.
	on one side and on the other two layers of $\frac{3}{4}$-inch Insulwood separated by 2-inch by $\frac{3}{4}$-inch fillets. Each set of studs on 14-inch centres	8·4	$7\frac{3}{4}$	36 (33)	49 (38)	78 (45)	54 (37)
73.	Eldorado slab cork with concrete-like surface on both sides	9·2	—	35*	34†	61‡	—
74.	A layer of $\frac{1}{4}$-inch L.W. brand fibre board on each side of 4-inch by 2-inch studding at 14-inch centres. Each face plastered to a thickness of $\frac{1}{2}$ inch	11·3	6	50 (36)	51 (40)	68 (46)	56 (40)
75.	A layer of $\frac{1}{2}$-inch Insulwood on each side of 4-inch by 2-inch studding at about 14-inch centres. One face plastered to a thickness of $\frac{3}{4}$ inch	11·6	$5\frac{3}{4}$	36 (36)	46 (40)	60 (46)	47 (40)
76.	A layer of $\frac{3}{4}$-inch Insulwood on each side of 4-inch by 2-inch studding at about 14-inch centres. One face plastered a thickness of $\frac{3}{4}$ inch	12·2	$6\frac{1}{4}$	36 (36)	46 (40)	66 (47)	49 (40)
77.	Staggered 4-inch by 2-inch stud-						

TABLE II. (*cont.*).

Partition number.	Description of test partition.			Average sound reduction for frequencies (cycles per second).			
	Construction.	Weight (lbs./sq. ft.).	Thickness (inches).	200 and 300.	500, 700, and 1000.	1600 and 2000.	200, 300, 500, 700, 1000, 1600, and 2000.
				db.	db.	db.	db.
	ding having a layer of ½-inch L.W. brand fibre board on each side. Each set of studs on 14-inch centres and each face plastered to a thickness of ½ inch.	12·2	7½	53 (36)	50 (40)	66 (47)	56 (40)
78.	Two sheets of plywood with the interspace filled with aluminium filings	14·0	3½	39*	28†	33§	—
79.	Lath and 3-coat lime plaster on each side of 4-inch by 2-inch studding on 12-inch centres....	16·4	6	38 (39)	45 (43)	63 (50)	48 (43)
80.	Two layers of Thatchboard, each 3 inches thick including facing of plaster on one side, insulated from one another. Plastered faces outward and a layer of Temperac quilt in the interspace..	16·6	—	40*	58†	75 §	—
81.	Lath and 3-coat hard plaster on each side of 4-inch by 2-inch studding on 12-inch centres....	17·0	5½	43 (39)	54 (43)	63 (50)	53 (43)
82.	Two leaves of Fosalsil blocks separated by ½-inch air-space. (Faces unplastered.)	18·3	4½	48*	45†	62§	—

TABLE II. (*cont.*).

Partition number.	Description of test partition. Construction.	Weight (lbs./sq. ft.).	Thickness (inches).	Average sound reduction for frequencies (cycles per second). 200 and 300.	500, 700, and 1000.	1600 and 2000.	200, 300, 500, 700, 1000, 1600, and 2000.
				db.	db.	db.	db.
83.	A layer of ½-inch Insulwood on each side of 4-inch by 2-inch studding at about 14-inch centres. Each face plastered to a thickness of ⅜ inch	18·8	6½	42 (40)	51 (44)	62 (50)	52 (44)
84.	Two leaves of Eonit 2-inch blocks separated by 1½-inch air-space, plastered on outer faces	18·9	6	47 (40)	45 (44)	58 (50)	49 (44)
85.	A layer of ¾-inch Insulwood on each side of 4-inch by 2-inch studding at about 14-inch centres. Each face plastered to a thickness of ¾ inch	19·3	7	44 (40)	52 (44)	67 (50)	54 (44)
86.	A layer of 1-inch Thermacoust wood fibre slabs on each side of 3-inch by 2-inch staggered studding at 16-inch centres. Each leaf and its studding insulated by ½-inch felt. Outer faces plastered to a thickness of ½ inch	21·9	9	52 (41)	54 (45)	66 (51)	57 (45)
87.	B.P. Bitumen laminated board $\frac{7}{16}$ inch thick, nailed to 1½-inch by 1-inch battens on one						

TABLE II. (*cont.*).

Partition number.	Description of test partition. Construction.	Weight (lbs./sq. ft.).	Thickness (inches).	Average sound reduction for frequencies (cycles per second). 200 and 300.	500, 700, and 1000.	1600 and 2000.	200, 300, 500, 700, 1000, 1600, and 2000.
				db.	db.	db.	db.
	side of 2-inch concrete. Board plastered on outer face to a thickness of $\frac{3}{8}$ inch	24·1	—	40■	44†	62‡	—
88.	Donnacona board on battens on one side of 2-inch concrete. Outer face of board plastered.	25·5	4	45*	48†	57‡	—
89.	A layer of Celotex on battens on one side of 2-inch concrete. Outer face of Celotex plastered.	26·7	—	46*	54†	54‡	—
90.	Two courses of Fosalsil blocks separated by $\frac{1}{2}$-inch air-space, outer surfaces plastered	27·4	$5\frac{3}{4}$	48*	52†	64§	—
91.	3-inch clinker slabs with fibre board stuck on each face; plastered on outer faces .	27·4	5	38 (43)	34 (47)	67 (53)	45 (48)
92.	2-inches of concrete on one side of a panel of Thatchboard 2 inches thick, plaster on the other side	28·1	—	45*	46†	59‡	—
93.	3-inch reinforced concrete with fibre board stuck on each face; outer faces plastered	45·0	5	46 (47)	36 (51)	73 (56)	50 (52)
94.	$4\frac{1}{2}$-inch brick with fibre board stuck on each face; outer faces plastered .	46·8	$6\frac{1}{4}$	44 (47)	46 (51)	73 (57)	53 (52)

the sound insulating value of a single homogeneous partition *.

It is found, however, that though agreement between measurement and calculation can be obtained for light partitions, the results for heavier panels such as brick walls fall below the calculated values, the divergences increasing with the weight. Excluding attempts to allow for panel resonances, the calculations predict that for the range of frequencies and weight covered by figs. 1–4 the insulation should increase fourfold (*i. e.*, by 6 db.) each time either the test frequency or the partition weight is doubled. The curves in figs. 1–4, which were obtained from the results given in Table I. A, show that this prediction is fulfilled for change of mass at low frequencies and for change of frequency with light partitions. In other words, for light simple partitions doubling the test frequency increases the insulation on an average by about 6 db., and for all simple partitions the insulation against low frequency sound is increased on an average by about 6 db. if the weight is doubled.

Berg † has recently reported that the results for wood panels tend to lie below the average for single panels of the same weight. This effect is shown by nos. 9 and 10, but not by the other two panels tested (nos. 7 and 12).

II. *Single Inhomogeneous Partitions.*

Comparison of the results for this group of partitions with the corresponding values for figs. 1 to 4 shows that the differences are only of the order of those occurring in the case of simple partitions. The only systematic deviation is a tendency for the results for hollow clay blocks and hollow tiles to be somewhat below the curve for high frequencies. It appears that the air-spaces contained in hollow blocks are of no acoustical value for insulating against air-borne sound, the only effect being a loss of insulation corresponding to the reduction in weight.

III. *Complex partitions.*

An important class of complex partition is that in which two layers of material are attached one on each side of a

* A. H. Davis, Phil. Mag. xv. p. 309 (1933) ; P. E. Sabine, J. Acous. Soc. America, iv. p. 38 (1932) ; R. Berger, ' Dissertation,' München (1911).

† Berg, *Det Kontelige Norske Videnskabers Selskab, Forhandlinger*, viii. no. 35 (1935).

wood framework. Usual building practice makes the framework of a number of 4-inch by 2-inch pieces of wood ("studs") spaced parallel to one another at about 14 to 18 inches between centres. Partitions numbered 61–65, 74, 79, 81, and 83 are examples of this. The sound reduction of this type of partition is almost invariably well above that of a single homogeneous partition having the same weight. It follows that the amplitude of the framework vibration is small (presumably owing to its rigidity), and the partition must be assumed to behave like several air-spaced partitions side by side. The result is of interest, as it shows that even with relatively stiff outside layers the mechanical coupling *via* the studding is small.

That there is little mechanical coupling between the outside layer and the studding is also shown by partitions numbered 58 and 59, in which fibre board is attached to one side of the studding only. The sound insulation of the partition is little greater than that of building board alone (see nos. 1–6) and is well below the mass curve, presumably because, although the studding has added considerably to the weight, the mechanical coupling between it and the fibre board is so small that the vibration of the latter is scarcely affected.

This class of partition must be distinguished from the staggered stud partition, in which the coupling between the two faces is minimized by attaching the two layers of building material to separate sets of studding. Partitions 60, 68, 77, and 86 are examples of this form of construction.

The results for partitions numbered 37 to 41 show that filling with light absorbent materials the space between two sheets of aluminium attached to a wood frame is of no advantage. Partition No. 57, which belongs to this series, contained a heavier filler, and it will be noted that the change of insulation with frequency is more similar to that of a simple partition. The measurements were, of course, confined to the air-borne sound insulation and do not indicate the reduction in drumming noises which might result from using the filler. The failure of the absorbent material to increase the sound insulation may be due to its mechanically linking the two faces of the partition. In some other experiments *, in which

* J. E. R. Constable, Proc. Phys. Soc. xlviii. p. 690 (1936).

such linkage was avoided, it has been shown that the sound insulation of a double partition can be increased by suitably disposing acoustical absorbent material within it.

Hurst * and Meyer † have shown that complex partitions which contain more than two leaves constitute an acoustical low-pass filter, the cut-off being sharper as more elements are used. One partition of this type (no. 43) has been measured, and it will be noted that although there is no clear evidence of a cut-off frequency there is a sharp rise of insulation between the middle and the high frequencies.

The results for the double windows, nos. 45–55, have already been published ‡, but are included here for ease of reference. They illustrate the effect of varying the spacing between two sheets of glass. In general the insulation increases with the spacing, but below a certain spacing for which the insulation is a minimum the reverse occurs. This effect is of importance chiefly at the lower frequencies, since the critical spacing is very small at high frequencies. The minimum in the insulation is due to the resonance of the oscillatory mechanical system formed by the masses of the sheets of glass and the stiffness of the air between.

A similar effect is obtained by interposing an elastic layer, *e. g.*, fibre board between a brick wall and its plaster facing (nos. 91, 93, and 94). In this case, however, the insulation at the upper frequencies is increased by the treatment, but a decrease of insulation in the middle and lower frequency ranges results in the reduction averaged for all frequencies showing no improvement. The results suggest that there is a resonance frequency round about 700∼. This would most simply be explained as due to the mechanical system formed by the masses of the wall and plaster coupled by the stiffness of the fibre board. This, however, necessitates the Young's modulus of the fibre board being about 100 times lower than its statical value. Possibly owing to the rough surface of the brick work the fibre board adheres to it at a relatively few points and hence its stiffness is considerably reduced.

* D. Hurst, Can. J. Research, xii. p. 398 (1935).
† E. Meyer, E. N. T. xii. p. 393 (1935).
‡ J. E. R. Constable, Phil. Mag. xviii. p. 321 (1934).

The authors express their gratitude to Dr. G. W. C. Kaye, Superintendent of the Physics Department, National Physical Laboratory, for his encouragement and advice, and to Dr. A. H. Davis for many helpful discussions. The authors also acknowledge their indebtedness to the many firms who have allowed publication of results obtained for their materials. Mr. M. de Selincourt obtained the results for partitions nos. 4, 22, 24, 28, 31, 33, 34, 43, 73, 80. 87, 89, and 92. Messrs. R. Berry and W. C. Copeland assisted with the measurements.

Summary.

Results of measurements made at the National Physical Laboratory of the sound insulating value of a number of partitions are given together with curves which exhibit the correlation between the sound insulation and weight of single homogeneous partitions. The results are discussed, and certain principles which should be adopted in sound insulating constructions are deduced.

XV. *The Carbon Arc in Vacuum.* *By* F. H. Newman, *D.Sc., F.Inst.P., Professor of Physics, University College, Exeter* *.

[Plate II.]

Various types of cold-cathode arcs in vacua have been previously described †, and the cathode current therein attributed to the emission of electrons from the cold cathode by the high field in the vicinity of this electrode. In all these cases the radiation is almost entirely that emitted from the residual gases within the arc tube—whatever these residual gases may be—and except in the case of the more volatile metals, such as sodium, mercury, etc., there is no indication of radiation corresponding to the vapour of the cathode metal. Thus the temperature of the cathode is far below that associated with, and generally attributed to, the ordinary carbon and tungsten arcs which operate on the thermionic principle. Druyvesteyn ‡ has suggested still another

* Communicated by the Author.
† Phil. Mag. xiv. p. 718 (1932).
‡ 'Nature,' cxxxvii. p. 580 (1936).

way in which the cathode of an arc may emit electrons, viz., where the electrode has very small areas covered with a thin layer of insulator; the positive ions on the latter give rise to a high field near the metal, and this field is strong enough to extract electrons from the cathode. On this view the cathode fall extends over the thickness of the insulator—probably of the order 10^{-4} or 10^{-5} cm.—and therefore produces the Langmuir "high field." In connexion with this suggestion it may be mentioned that the author has referred in the previous papers to the fact that an arc in a vacuum can be started and maintained much more easily if there are particles of foreign materials, such as oxides, on the cathode, and this result had been attributed to such particles forming the sharp or submicroscopic points necessary for the "high field" theory. At these points the electrical force is greater than that existing in their immediate neighbourhood, and the electron currents drawn from these points are much greater than are to be expected on any theory.

This "high field" theory of the arc is not accepted by many investigators who stand by the thermionic theory, although measurements of the cathode temperature, even in the same type of arc, are very discordant, and by no means justify this theory. It is difficult to substantiate either theory from measurements made with any arc after it has been burning for such a time-interval that the electrodes are incandescent. Even in the case of the so-called cold arcs there is the cathode bright spot, the temperature of which may be high enough to cause the requisite copious emission of electrons necessary to account for the cathode current. It has seemed to the author that the most useful information is likely to be obtained from a study of the starting of an arc, especially if the arc is struck, not in the usual way by drawing the electrodes apart, but by means of an electrical discharge between separated arc electrodes across which a potential difference is applied. The technique of this method has been developed, but it is only successful when the arc operates at low pressure, *i. e.*, in a vacuum. All metallic arcs can be started and maintained in this manner, and, apart from its convenience, *i. e.*, low pressures giving sharp spectrum lines and little, if any, self-reversal, the method throws much light on the mechanism of the commencement of an arc. It cannot

be expected to provide much information about the mechanism of an arc which has been burning for so long that the electrodes are incandescent, or, if the bright cathode spot is present, because during maintenance the mechanism of such arcs does not differ from that of arcs started in the usual manner. In the latter type, where the electrodes are first in contact and then drawn apart, at the instant and place at which contact is broken there must be a localized high temperature with consequent thermionic emission. This emission cannot possibly occur with the starting device used by the author in high vacuum arcs.

The present paper describes experiments with the cold-cathode arc between graphite electrodes in gases at very low pressures. The graphite employed was in the form of rods 6·3 mm. thick and of special purity for spectrographic purposes, the degree of purity being 99·98 per cent. Spectrographic analysis gave the following lines in the visible region from the impurities :—calcium, λ 4226·7, sodium D lines, vanadium λ 4407·9 and lithium λ 6707·9. This pure graphite was selected because, as previously mentioned, it had been found that the existence of impurities on the cathode facilitated the striking of high-vacuum arcs. Although it has been suggested that such impurities may form the sharp points necessary for the high-field theory, there is the possibility that thermionic emission takes place from the impurities when the electrodes are bombarded by the ions, and it was desired to eliminate this effect as far as possible.

The arc tube was of pyrex, with tungsten rods sealed into the glass. Small holes bored along the graphite electrodes enabled the latter to be fixed to the tungsten, and this arrangement eliminated wax seals, which are always unsatisfactory where a considerable current is passing. In the first series of experiments a third graphite electrode was inserted in the tube, and between the first two electrodes, the ends of which were separated 5 mm., a steady potential difference of 100 volts was applied. Any voltage above 70 sufficed, but the higher the applied potential difference the more easily was the arc started and maintained. Actually the arc was started, as in all previous experiments, by passing a momentary electrical discharge between one of the two arc electrodes—whether the negative or positive one was immaterial—and the

third distant electrode. The gas-pressure necessary for successful operation ranged from 1 to 0·01 mm. of mercury, and the arc-current throughout the experiments was maintained at 5 amperes.

Under these conditions the arc could be started and maintained whether the gas present was air or hydrogen, although the arc was struck more readily with hydrogen. Owing to the liberation of gas from the carbon electrodes it was necessary to keep the pump running continuously, while with hydrogen it was arranged to have a continuous flow of the gas into the arc-tube throughout the experiment. The radiation from the arc at the moment of starting was photographed, an exposure of 3 seconds being sufficient with the intense light. In this short period the carbon electrodes did not reach a high temperature, so there was no question of thermionic emission. Under these circumstances one would not expect the typical carbon-arc spectrum that is obtained with graphite electrodes in air at atmospheric pressure, and this expectation was realized, although the CN bands appear. The two spectra are shown in Pl. II. (I. and II.), and it is evident that the radiation from the arc in a vacuum at the moment of starting comes from the ionized residual gases in the tube. The D lines of sodium in Pl. II. (II.) are prominent together with the CN bands λ 4216; the Swan bands also appear, but both band systems are attributed to the action of the arc discharge on the mixture of carbon dioxide, carbon monoxide, nitrogen, and hydrogen present in the tube, rather than to a thermal reaction between nitrogen and hydrogen on the hot carbon, or carbon vapour.

With hydrogen, lines of the Balmer series and the secondary spectrum of hydrogen are strongly marked, Pl. II. (III.); the Swan bands, together with the CH band at λ 4315, are also prominent. For comparison, a spectrogram, Pl. II. (IV.), of the electrical discharge through a hydrogen vacuum tube is shown. The similarity of these two spectrograms and the difference between I. and II. indicate that if thermionic emission is responsible for the arc maintenance in I., it does not enter into the mechanism of the striking of arcs in vacua. The graphite rods contained considerable quantities of occluded hydrogen, and if the arc was allowed to burn in the residual air, cut off, and then restarted, hydrogen

lines from the gas liberated from the graphite were prominent together with the CH band, λ 4315, as shown in Pl. II. (V.).

In the next series of experiments an electrical discharge between two aluminium electrodes was passed perpendicularly across the space between the two graphite electrodes —still separated by 5 mm.—across which the 100 volts direct potential difference was applied. Under these conditions it is reasonable to assume that no high-velocity electrons or ions strike against either of the graphite electrodes, so that there can be no localized heating at any point thereon. In other words, any thermionic effect at the graphite is eliminated. But the arc could be readily started by passing a momentary electrical discharge between the two aluminium electrodes. In this case the starting of the arc is attributed to the fact that the positive ions, produced by the electrical discharge, form the space-charge of positive ions in the immediate neighbourhood of the negative graphite electrode, and so give rise to the high field at the cathode. This type of arc could be started and maintained both in air and hydrogen at low pressures, and the spectra of the emitted radiation were similar to those described above and shown in Pl. II.

In a third series of experiments two graphite arc electrodes alone were used in the tube, and an electrodeless discharge employed to produce the momentary electrical impulse. Six turns of heavy copper wire were wound round the outside of the tube so as to enclose the graphite electrodes, and this coil was connected to Leyden jars and an induction coil in the usual manner. A momentary electrodeless discharge started the arc between the graphite electrodes, the latter still being 5 mm. apart and connected to 100 volts D.C. It was, however, more difficult to strike the arc under these conditions, a powerful discharge being required, and hydrogen gas was more effective than air. The gas-pressure was critical—0·01 mm. of mercury. Again there is no question of thermionic action, the effect being due to the positive ion space-charge at the cathode.

In all these experiments it should be emphasized that the electrical discharge was only passed for a very short time interval—less than one second—and no continuous discharge was required. On the contrary, if for any reason the arc did not start, a continued passage of the discharge was not effective. This fact further strengthens

the view that there is no thermal effect, and that the arc is started by the rapid surge of positive ions to form the positive ion space-charge at the cathode.

XVI. *The Characteristic Property of Saint-Venant's Solutions for the Torsion and Bending of an Elastic Cylinder* *. *By* J. N. Goodier, *M.A., Ph.D.*

SAINT-VENANT was led to his well-known solutions for the problems of torsion and bending in elastic rods by a set of physically plausible assumptions. These were equivalent to

$$X_x = X_y = Y_y = Z_z = 0$$

for torsion (Oz being parallel to the generators of the cylindrical surface); and for cantilever bending to

$$X_x = X_y = Y_y = 0 \quad \text{and} \quad Z_z = (c - z)(Ax + By + C),$$

the four constants occurring being determined by the resultant actions on the ends and the choice of origin.

Voigt [1] showed that the assumption that all six components of stress are independent of z leads to (i.) pure tension, (ii.) flexure by couples, and (iii.) torsion (of Saint-Venant's types), and that the assumption that all six components are linear functions of z leads in addition to (iv.) cantilever bending (of Saint-Venant's type). Clebsch [1] showed that the assumption $X_x = X_y = Y_y = 0$, *i. e.*, no normal traction across any plane parallel to a generator, leads to all four (Saint-Venant) types of deformation. These results constitute some simplification of Saint-Venant's assumptions. This paper establishes further simplification. It will be proved that Saint-Venant's solutions follow from the following postulate :—

> *The three components of stress* (X_z, Y_z, Z_z) *acting on a cross-section* (z = *constant*) *are propagated along the rod according to the same law as the resultants they constitute* . . . . . (A)

An alternative form, perhaps more satisfactory from the physical view-point, is :—

> *The distribution on the cross-section of the stress constituting each component of the resultant force and couple is the same at all sections.* . . . . . (B)

* Communicated by the Author.

Considering first tension, torsion, and flexure (by couples only), the resultants consist of an axial force and torsional and flexural couples, which are propagated along the bar without change. The stresses which constitute them are X_z, Y_z, and Z_z, and our initial assumptions (which are sufficient to define the Saint-Venant solutions) are therefore

$$\frac{\partial X_z}{\partial z} = \frac{\partial Y_z}{\partial z} = \frac{\partial Z_z}{\partial z} = 0. \quad . \quad . \quad . \quad . \quad (1)$$

These follow also from (B).

Secondly, when there are transverse shearing resultants (cantilever bending), the flexural moments become linear functions of z. The stress component constituting these moments is Z_z. The resultants which X_z, Y_z constitute are still propagated without change. The initial assumptions are then

$$\frac{\partial X_z}{\partial z} = \frac{\partial Y_z}{\partial z} = \frac{\partial^2 Z_z}{\partial z^2} = 0, \quad . \quad . \quad . \quad . \quad (2)$$

which also follow from B, and again these are sufficient to define the Saint-Venant solutions. But only the first two are necessary, the third then following from the last of (3).

Torsion, Flexure, and Tension.

The general equations of equilibrium and compatibility are

$$\left.\begin{aligned} &\frac{\partial X_x}{\partial x} + \frac{\partial X_y}{\partial y} + \frac{\partial X_z}{\partial z} = 0, \quad \frac{\partial Y_y}{\partial y} + \frac{\partial Y_z}{\partial z} + \frac{\partial X_y}{\partial x} = 0, \\ &\frac{\partial Z_z}{\partial z} + \frac{\partial X_z}{\partial x} + \frac{\partial Y_z}{\partial y} = 0, \end{aligned}\right\} \quad (3)$$

$$\left.\begin{aligned} &(1+\sigma)\nabla^2 X_x + \frac{\partial^2 \Theta}{\partial x^2} = 0, \quad (1+\sigma)\nabla^2 Y_y + \frac{\partial^2 \Theta}{\partial y^2} = 0, \\ &(1+\sigma)\nabla^2 Z_z + \frac{\partial^2 \Theta}{\partial z^2} = 0, \end{aligned}\right\} \quad (4)$$

$$\left.\begin{aligned} &(1+\sigma)\nabla^2 X_y + \frac{\partial^2 \Theta}{\partial x\,\partial y} = 0, \quad (1+\sigma)\nabla^2 Y_z + \frac{\partial^2 \Theta}{\partial y\,\partial z} = 0, \\ &(1+\sigma)\nabla^2 X_z + \frac{\partial^2 \Theta}{\partial x\,\partial z} = 0, \end{aligned}\right\} \quad (5)$$

where $\Theta = X_x + Y_y + Z_z$, $\nabla^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$, and σ is Poisson's ratio.

In view of (1), which implies, of course, that $\frac{\partial^2}{\partial z^2}(X_z, Y_z, Z_z)$ vanish, the equations become

$$\frac{\partial X_x}{\partial x} + \frac{\partial X_y}{\partial y} = 0, \quad \frac{\partial Y_y}{\partial y} + \frac{\partial X_y}{\partial x} = 0, \quad \frac{\partial X_z}{\partial x} + \frac{\partial Y_z}{\partial y} = 0, \quad (6)$$

$$\left.\begin{aligned} (1+\sigma)\nabla^2 X_x + \frac{\partial^2 \Theta}{\partial x^2} = 0, \quad (1+\sigma)\nabla^2 Y_y + \frac{\partial^2 \Theta}{\partial y^2} = 0, \\ (1+\sigma)\nabla_1^2 Z_z + \frac{\partial^2}{\partial z^2}(X_x + Y_y) = 0, \end{aligned}\right\} \quad (7)$$

$$\left.\begin{aligned} (1+\sigma)\nabla^2 X_y + \frac{\partial^2 \Theta}{\partial x \partial y} = 0, \quad (1+\sigma)\nabla_1^2 Y_z + \frac{\partial^2 \Theta}{\partial y \partial z} = 0, \\ (1+\sigma)\nabla_1^2 X_z + \frac{\partial^2 \Theta}{\partial x \partial z} = 0, \end{aligned}\right\} \quad (8)$$

where $$\nabla_1^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}.$$

Now $\nabla^2 \Theta = 0$ by addition of (4), and therefore

$$\nabla_1^2 (X_x + Y_y + Z_z) + \frac{\partial^2}{\partial z^2}(X_x + Y_y) = 0. \quad . \quad . \quad (9)$$

Differentiating the third of (7) with respect to z, we have

$\frac{\partial^3}{\partial z^3}(X_x + Y_y) = 0$, and then, from (9)

$$\nabla_1^2 \left(\frac{\partial X_z}{\partial z} + \frac{\partial Y_y}{\partial z} \right) = 0. \quad . \quad . \quad . \quad . \quad (10)$$

If now we differentiate the first two of (6) with respect to z, we have, with (10), three equations in $\frac{\partial X_x}{\partial z}$, $\frac{\partial X_y}{\partial z}$, $\frac{\partial Y_y}{\partial z}$, and *these equations have the same form as the differential equations of plane strain.* They are satisfied if

$$\frac{\partial X_x}{\partial z} = \frac{\partial^2 \chi}{\partial y^2}, \quad \frac{\partial Y_y}{\partial z} = \frac{\partial^2 \chi}{\partial x^2}, \quad \frac{\partial X_y}{\partial z} = -\frac{\partial^2 \chi}{\partial x \partial y},$$

provided $\nabla_1^4\chi=0$. Thus $\frac{\partial X_x}{\partial z}$, $\frac{\partial X_y}{\partial z}$, $\frac{\partial Y_y}{\partial z}$ at any cross-section are determined by the z-derivatives of the boundary stress at that section in the same way as the stresses in plane strain are determined by the boundary stress. The stress on the curved surface of the rod, and its z-derivatives, vanish. We may therefore conclude without further analysis that $\frac{\partial X_x}{\partial z}$, $\frac{\partial X_y}{\partial z}$, and $\frac{\partial Y_y}{\partial z}$ vanish all over the section, dislocational stress being, of course, specifically excluded if the section is not simply connected.

It follows that $\frac{\partial^2}{\partial z^2}$ (X_x, X_y, Y_y) vanish also, and then from (7) $\nabla_1^2 Z_z=0$, so that, from (9)

$$\nabla_1^2(X_x+Y_y)=0.$$

This and the first two of (6) are again the equations of plane strain, this time in X_x, X_y, Y_y. Since the stress on the boundary of the section vanishes, they show that X_x, X_y, Y_y vanish. We now find from (4) and the first of (5) that $\left(\frac{\partial^2}{\partial x^2}, \frac{\partial^2}{\partial x \partial y}, \frac{\partial^2}{\partial y^2}\right)Z_z=0$, or that Z_z is linear in x and y, and by hypothesis it is independent of z; Z_z constitutes the tension and the two components of the flexural couple, so that it is completely determined when these resultants are given. Thus the tension and flexure are independent of each other and of the torsion.

The outstanding equations, the third of (6) and the last two of (8), become

$$\frac{\partial X_z}{\partial x}+\frac{\partial Y_z}{\partial y}=0, \quad \nabla_1^2 X_z=0, \quad \nabla_1^2 Y_z=0,$$

and these lead unambiguously to Saint-Venant's solutions [2] for torsion.

Cantilever Bending.

With the assumptions (2), the equations (3) and (4) become

$$\frac{\partial X_x}{\partial x}+\frac{\partial X_y}{\partial y}=0, \quad \frac{\partial Y_y}{\partial y}+\frac{\partial X_y}{\partial x}=0, \quad \frac{\partial X_z}{\partial x}+\frac{\partial Y_z}{\partial y}+\frac{\partial Z_z}{\partial z}=0, \quad \ldots \quad (11)$$

$$\left.\begin{aligned}(1+\sigma)\nabla^2 X_x+\frac{\partial^2\Theta}{\partial x^2}=0,\quad (1+\sigma)\nabla^2 Y_y+\frac{\partial^2\Theta}{\partial y^2}=0,\\ (1+\sigma)\nabla_1^2 Z_z+\frac{\partial^2}{\partial z^2}(X_x+Y_y)=0.\end{aligned}\right\}\quad (12)$$

Equation (9) follows from $\nabla^2\Theta=0$ as before. By differentiation of the last of (12) twice with respect to z, we have $\frac{\partial^4}{\partial z^4}(X_x+Y_y)=0$, and (9) then yields

$$\nabla_1^2\left(\frac{\partial^2 X_x}{\partial z^2}+\frac{\partial^2 Y_y}{\partial z^2}\right)=0.$$

The first two of (11), differentiated twice with respect to z, form with this equation the equations of plane strain with $\frac{\partial^2 X_x}{\partial z^2}$, $\frac{\partial^2 X_y}{\partial z^2}$, $\frac{\partial^2 Y_y}{\partial z^2}$ playing the parts of the stresses. By the same argument as before we deduce that these derivatives vanish, then that

$$\nabla_1^2(X_x+Y_y)=0,$$

and hence that X_x, X_y, Y_y vanish everywhere. Again Z_z is a linear function of x and y. By hypothesis $\frac{\partial^2 Z_z}{\partial z^2}$ vanishes, and therefore

$$Z_z=A+Bx+Cy+z(D+Ex+Fy),$$

the arbitrary constants being determined by the resultants and the choice of origin.

The outstanding equations are, from (11), (12), and (5)

$$\frac{\partial X_z}{\partial x}+\frac{\partial Y_z}{\partial y}=-\frac{\partial Z_z}{\partial z},\quad (1+\sigma)\nabla_1^2 X_z=-\frac{\partial^2 Z_z}{\partial x\partial z},$$

$$(1+\sigma)\nabla_1^2 Y_z=-\frac{\partial^2 Z_z}{\partial y\partial z},$$

and, since Z_z is already independently determined, they suffice to determine X_z, Y_z in conjunction with the boundary conditions. They lead directly to the Saint-Venant solutions [3] for cantilever bending. The bending is associated with torsion unless the transverse force resultants pass through the "centre of twist" of the cross-section.

References.

(1) *Cf.* (i.) Love, 'Mathematical Theory of Elasticity,' 1927, p. 21. (ii.) Southwell, Proc. Roy. Soc. A, 881, cliv. p. 4 (1936).
(2) Timoshenko, 'Theory of Elasticity,' 1934, p. 230.
(3) *Ibid.* p. 286.

Ontario Research Foundation.

XVII. *Notices respecting New Books.*

An Introduction to the Theory of Elasticity for Engineers and Physicists. By Professor R. V. Southwell. (Pp. 518, with two plates and many diagrams.] (Oxford : University Press. Price 30*s.* net.)

THE author states that he has attempted, while covering the usual ground of an honours course in Strength of Materials, to provide the reader with the "background" needed for more advanced study and research. With this aim he has endeavoured, first, to lay stress on the physical principles which underlie the whole theory of Elasticity, and, secondly, to indicate the development of mathematical technique from simple solutions, based on appeal to intuition, up to the stage at which the subject acquires an interest for the professional mathematician. The author also hopes that, after reading Chapters VIII. and XII. (which are of an advanced mathematical character), the student will be fitted to take profit from Professor Love's much greater work ; and from the chapters devoted to elastic stability and vibrations he should acquire some grasp of the powerful methods which we owe to the late Lord Rayleigh.

Discussion of the experimental side of the subject (of Strength of Materials), often called the Testing of Materials, although not ignored, is not a primary object of the work (reviewer's remark).

The objects of the author appear to have been admirably carried out, and the result is a lucid, reasoned volume. The value of the book is enhanced by copious examples taken (by permission) from examination papers set at Cambridge and Oxford. Thorough acquaintance with theoretical engineering principles can only come by constant practice with problems involving calculations, and the examples referred to are most instructive.

Professor Southwell's method of commencing his subject (by discussing such subjects as The Principle of Superposition, Elastic Strain Energy, the Reciprocal Theorem, and Castigliano's Theorems) is unusual, but eminently successful. Incidentally, the importance of Reciprocal Theorems in Physics generally is well illustrated by the use of such Theorems in Elasticity. Professor Southwell's footnote (p. 96) on the subject of Castigliano's papers is strongly reminiscent of footnotes that had to be made years ago about Osborne Reynolds.

A few other points may be mentioned. The author explains the basis of the important "Moment Distribution Method" of Professor Hardy Cross for the treatment of frameworks having rigid joints, but regards a detailed description as outside the scope of his work. The soap film method of

dealing with the torsion of non-circular shafts, and other allied matters, is also dealt with.

Rayleigh's method of dealing with problems involving either vibrations or stability of elastic systems has already been referred to.

Generally speaking, the second half of the book is more mathematical and therefore harder to read than the first half. It is just possible that this may daunt some students from studying such subjects as stresses in long circular shafts, rotating disks, thick cylinders, etc. This, however, is a minor criticism.

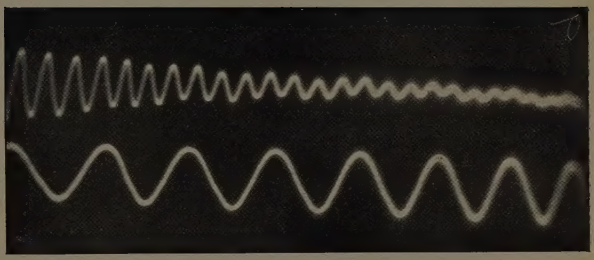
The book is well produced by the Oxford University Press, and there are copious footnotes and references. Curiously enough, Greenhill's name is not mentioned in connexion with the wire-wound gun.

The book is a welcome addition to the literature of the subject.

Simplified Structure Factor and Electron Density Formulæ for the 230 *Space Groups of Mathematical Crystallography.* By Dr. KATHLEEN LONSDALE. (Published for the Royal Institution.) [Pp. vii+181.] (London: G. Bell and Sons, Ltd., 1936. Price 10*s.* net.)

THE space group tables produced twelve years ago by W. T. Astbury and Mrs. Lonsdale (then Miss Yardley) have been of the greatest service to all workers engaged in the analysis of crystal structures. Mrs. Lonsdale has now increased this indebtedness by working out the expressions for the structure factors for the 230 space groups. Anyone who has calculated from first principles the structure factors for a few space groups and attempted to express the results in the most concise and useful form will appreciate the results of this heroic task which Mrs. Lonsdale has carried out. To quote from the publishers' announcement: "In order to reduce the cost to the reader and also to minimise the possibility of error, the author's own script has been reproduced by a photographic method." The beautiful clearness of the handwriting justifies this method of producing the book, and if authors generally could be persuaded to write as clearly, this method of producing scientific works might be more frequently used. In addition to the tables, there is a useful introduction which explains very clearly the use of structure factors in the analysis of crystals and in the calculation of electron density distributions. Finally, it only remains to say that the general appearance of the book is very pleasing and that it can be unreservedly recommended to all workers in this sphere of research.

Fig. (*a*)



Time-scale 50 per second.

Fig. (*b*)

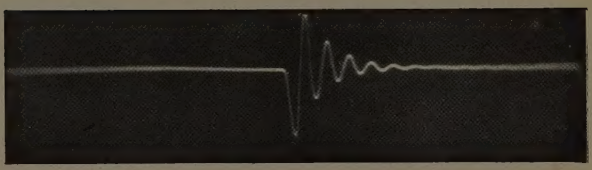
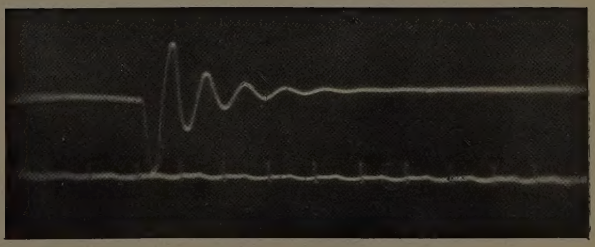
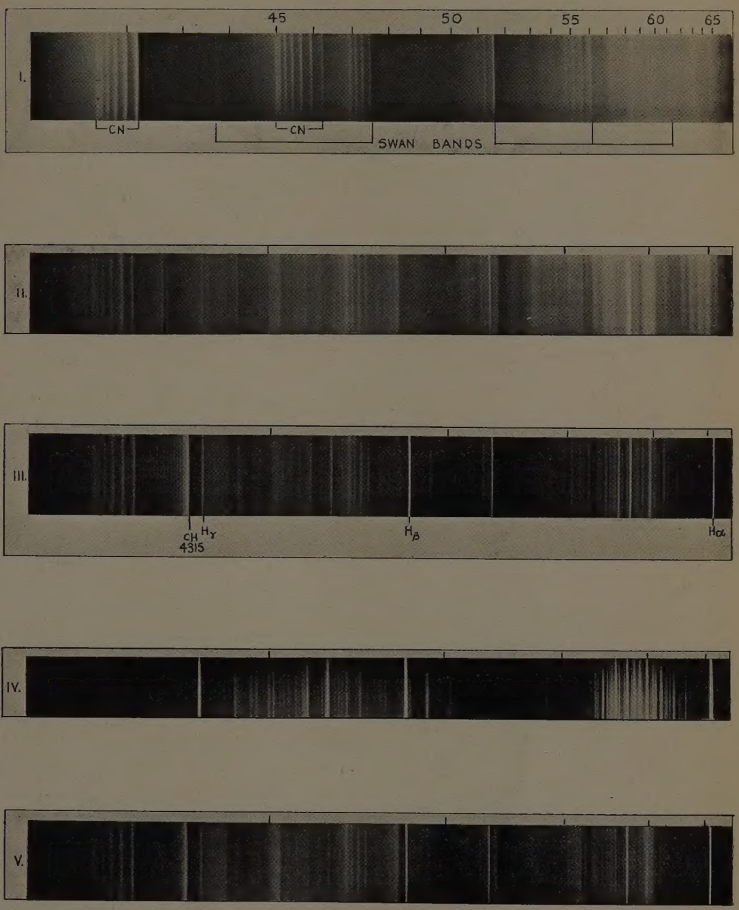


Fig. (*c*)



Time-scale for figs. (*b*) and (*c*) 1024 per second.

Records of flexural vibration: fig. (*a*) wallboard; fig. (*b*) rock asphalt; fig. (*c*) asphalt mastic.



The carbon arc.

I. Arc in atmosphere.
II. Arc in vacuum (residual gases).
III. Arc in vacuum (hydrogen).
IV. Electric discharge-tube (hydrogen),
V. Arc in vacuum (residual gases after arc has been burning for period).